

光子学研究

ISSN 2327-9125 CN 31-2126/O4

# PHOTONICS Research

Volume 9 | 2021

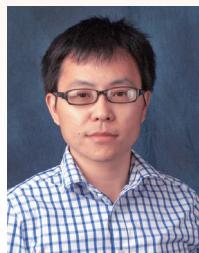
## Special Issue on

# DEEP LEARNING IN PHOTONICS



# GUEST EDITORS

*Special Issue / Deep Learning in Photonics*



Zongfu Yu

University of Wisconsin, Madison, USA



Yang Chai

The Hong Kong Polytechnic University, China



Li Gao

Nanjing University of Posts and Telecommunications, China



Darko Zibar

Technical University of Denmark, Denmark



# CONTENTS

**PHOTONICS  
Research**

Special Issue on  
Deep Learning in Photonics

Deep learning in photonics: introduction

*Li Gao, Yang Chai, Darko Zibar, and Zongfu Yu*

**1**

*Photon. Res. 9(8), DLP1-DLP3 (2021)*

Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing

*Qiang Cai, Ya Guo, Pu Li, Adonis Bogris, K. Alan Shore, Yamei Zhang, and Yuncai Wang*

**4**

*Photon. Res. 9(1), B1-B8 (2021)*

Integrating deep learning to achieve phase compensation for free-space orbital-angular-momentum-encoded quantum key distribution under atmospheric turbulence

*Xingyu Wang, Tianyi Wu, Chen Dong, Haonan Zhu, Zhuodan Zhu, and Shanghong Zhao*

**12**

*Photon. Res. 9(2), B9-B17 (2021)*

Deep plug-and-play priors for spectral snapshot compressive imaging

*Siming Zheng, Yang Liu, Ziyi Meng, Mu Qiao, Zhishen Tong, Xiaoyu Yang, Shensheng Han, and Xin Yuan*

**21**

*Photon. Res. 9(2), B18-B29 (2021)*

High-fidelity image reconstruction for compressed ultrafast photography via an augmented-Lagrangian and deep-learning hybrid algorithm

*Chengshuai Yang, Yunhua Yao, Chengzhi Jin, Dalong Qi, Fengyan Cao, Yilin He, Jiali Yao, Pengpeng Ding, Liang Gao, Tianqing Jia, Jinyang Liang, Zhenrong Sun, and Shian Zhang*

**33**

*Photon. Res. 9(2), B30-B37 (2021)*

Smart ring resonator-based sensor for multicomponent chemical analysis via machine learning

*Zhenyu Li, Hui Zhang, Binh Thi Thanh Nguyen, Shaobo Luo, Patricia Yang Liu, Jun Zou, Yuzhi Shi, Hong Cai, Zhenchuan Yang, Yufeng Jin, Yilong Hao, Yi Zhang, and Ai-Qun Liu*

**41**

*Photon. Res. 9(2), B38-B44 (2021)*

Monte Carlo simulation fused with target distribution modeling via deep reinforcement learning for automatic high-efficiency photon distribution estimation

*Jianhui Ma, Zun Piao, Shuang Huang, Xiaoman Duan, Genggeng Qin, Linghong Zhou, and Yuan Xu*

**48**

*Photon. Res. 9(3), B45-B56 (2021)*

Deep compressed imaging via optimized pattern scanning [On the Cover]

*Kangning Zhang, Junjie Hu, and Weijian Yang*

**60**

*Photon. Res. 9(3), B57-B70 (2021)*

Backpropagation through nonlinear units for the all-optical training of neural networks

*Xianxin Guo, Thomas D. Barrett, Zhiming M. Wang, and A. I. Lvovsky*

**74**

*Photon. Res. 9(3), B71-B80 (2021)*

---

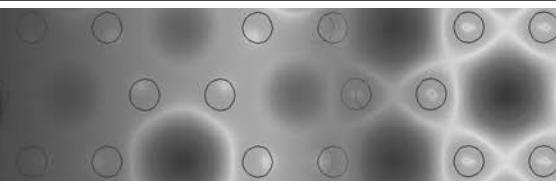
## On the Cover

A new deep compressed imaging modality enables high speed image acquisition and high fidelity object reconstruction.

Learning to recognize misaligned hyperfine orbital angular momentum modes	Xiao Wang, Yufeng Qian, JingJing Zhang, Guangdong Ma, Shupeng Zhao, RuiFeng Liu, Hongrong Li, Pei Zhang, Hong Gao, Feng Huang, and Fuli Li <i>Photon. Res.</i> 9(4), B81-B86 (2021)	84
Experimental study of neuromorphic node based on a multiwaveband emitting two-section quantum dot laser	George Sarantoglou, Menelaos Skontransis, Adonis Bogris, and Charis Mesaritakis <i>Photon. Res.</i> 9(4), B87-B95 (2021)	90
Engineering of multiple bound states in the continuum by latent representation of freeform structures	Ronghui Lin, Zahrah Alnakhli, and Xiaohang Li <i>Photon. Res.</i> 9(4), B96-B103 (2021)	99
Real-time deep learning design tool for far-field radiation profile	Jinran Qie, Erfan Khoram, Dianjing Liu, Ming Zhou, and Li Gao <i>Photon. Res.</i> 9(4), B104-B108 (2021)	107
Sensing in the presence of strong noise by deep learning of dynamic multimode fiber interference	Linh V. Nguyen, Cuong C. Nguyen, Gustavo Carneiro, Heike Ebendorff-Heidepriem, and Stephen C. Warren-Smith <i>Photon. Res.</i> 9(4), B109-B118 (2021)	112
Delay-weight plasticity-based supervised learning in optical spiking neural networks	Yanan Han, Shuiying Xiang, Zhenxing Ren, Chentao Fu, Aijun Wen, and Yue Hao <i>Photon. Res.</i> 9(4), B119-B127 (2021)	122
Free-space optical neural network based on thermal atomic nonlinearity	Albert Ryou, James Whitehead, Maksym Zhelyeznyakov, Paul Anderson, Cem Keskin, Michal Bajcsy, and Arka Majumdar <i>Photon. Res.</i> 9(4), B128-B134 (2021)	131
Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks	Yihao Xu, Xianzhe Zhang, Yun Fu, and Yongmin Liu <i>Photon. Res.</i> 9(4), B135-B152 (2021)	138
On-demand design of spectrally sensitive multiband absorbers using an artificial neural network	Sunae So, Younghwan Yang, Taejun Lee, and Junsuk Rho <i>Photon. Res.</i> 9(4), B153-B158 (2021)	156
Intelligent coding metasurface holograms by physics-assisted unsupervised generative adversarial network	Che Liu, Wen Ming Yu, Qian Ma, Lianlin Li, and Tie Jun Cui <i>Photon. Res.</i> 9(4), B159-B167 (2021)	162
Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images	Zafraan Hussain Shah, Marcel Müller, Tung-Cheng Wang, Philip Maurice Scheidig, Axel Schneider, Mark Schüttelpelz, Thomas Huser and Wolfram Schenck <i>Photon. Res.</i> 9(5), B168-B181 (2021)	171
Deep learning in nano-photonics: inverse design and beyond	Peter R. Wiecha, Arnaud Arbouet, Christian Girard, and Otto L. Muskens <i>Photon. Res.</i> 9(5), B182-B200 (2021)	185

All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes	<i>Yahui Zhang, Joshua Robertson, Shuiying Xiang, Matěj Hejda, Julián Bueno, and Antonio Hurtado</i>	<b>204</b>
Imaging through unknown scattering media based on physics-informed learning	<i>Shuo Zhu, Enlai Guo, Jie Gu, Lianfa Bai, and Jing Han</i>	<b>213</b>
Photon. Res. 9(5), B201-B209 (2021)		
Incoherent imaging through highly nonstatic and optically thick turbid media based on neural network	<i>Shanshan Zheng, Hao Wang, Shi Dong, Fei Wang, and Guohai Situ</i>	<b>223</b>
Photon. Res. 9(5), B220-B228 (2021)		
Realizing transmitted metasurface cloak by a tandem neural network	<i>Zheng Zhen, Chao Qian, Yuetian Jia, Zhixiang Fan, Ran Hao, Tong Cai, Bin Zheng, Hongsheng Chen, and Erping Li</i>	<b>232</b>
Photon. Res. 9(5), B229-B235 (2021)		
Accurate inverse design of Fabry–Perot-cavity-based color filters far beyond sRGB via a bidirectional artificial neural network	<i>Peng Dai, Yasi Wang, Yueqiang Hu, C. H. de Groot, Otto Muskens, Huigao Duan, and Ruomeng Huang</i>	<b>239</b>
Photon. Res. 9(5), B236-B246 (2021)		
Genetic-algorithm-based deep neural networks for highly efficient photonic device design	<i>Yangming Ren, Lingxuan Zhang, Weiqiang Wang, Xinyu Wang, Yufang Lei, Yulong Xue, Xiaochen Sun, and Wenfu Zhang</i>	<b>250</b>
Photon. Res. 9(6), B247-B252 (2021)		
Toward simple, generalizable neural networks with universal training for low-SWaP hybrid vision	<i>Baurzhan Muminov, Altai Perry, Rakib Hyder, M. Salman Asif, and Luat T. Vuong</i>	<b>256</b>
Photon. Res. 9(7), B253-B261 (2021)		
Towards smart optical focusing: deep learning-empowered dynamic wavefront shaping through nonstationary scattering media	<i>Yunqi Luo, Suxia Yan, Huanhao Li, Puxiang Lai, and Yuanjin Zheng</i>	<b>265</b>
Photon. Res. 9(8), B262-B278 (2021)		

# PHOTONICS Research



## Deep learning in photonics: introduction

LI GAO,<sup>1,5</sup> YANG CHAI,<sup>2,6</sup> DARKO ZIBAR,<sup>3,7</sup> AND ZONGFU YU<sup>4,8</sup>

<sup>1</sup>Institute of Advanced Materials (IAM), and School of Materials Science and Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210046, China

<sup>2</sup>Department of Applied Physics, The Hong Kong Polytechnic University, Hong Kong, China

<sup>3</sup>Department of Photonics Engineering, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark

<sup>4</sup>Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

<sup>5</sup>e-mail: iamlgao@njupt.edu.cn

<sup>6</sup>e-mail: ychai@polyu.edu.hk

<sup>7</sup>e-mail: dazi@fotonik.dtu.dk

<sup>8</sup>e-mail: zyu54@wisc.edu

Received 22 April 2021; posted 22 April 2021 (Doc. ID 428702); published 19 July 2021

The connection between Maxwell's equations and neural networks opens unprecedented opportunities at the interface between photonics and deep learning. This feature issue highlights recent research progress at the interdisciplinary field of photonics and deep learning and provides an opportunity for different communities to exchange their ideas from different perspectives. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.428702>

The application of deep learning in photonics has gained a tremendous amount of attention in the past few years. This interdisciplinary research covers a broad range of topics, including the inverse design of photonic devices, enhanced sensing and imaging, neuromorphic computing, and many other emerging applications. This feature issue provides a snapshot of current research with a collection of recent advances in this thriving field. We feature 28 papers contributed by experts from various aspects of photonics research.

The first collection of works in this issue focuses on using deep learning for device design. Conventional design of photonic structures often requires solving Maxwell's equations on a large scale. However, our capability to solve Maxwell's equation has significantly lagged behind today's fabrication capability that routinely produces devices with millions of nanostructure features. Using neural networks for photonic design allows researchers to tap into a rich set of machine-learning algorithms which enable fast inverse design and optimization. Examples include on-demand design of spectrally sensitive multi-band absorber [1], Fabry-Pérot-cavity-based color filters far beyond sRGB [2], transmitted metasurface cloak by tandem neural network [3] and intelligent coding metasurface hologram by physics assisted unsupervised generative adversarial network [4]. Strategy that combines deep neural network with genetic algorithms requires significantly less training data and enables highly efficient inverse designs [5]. For a more comprehensive review of past research and future perspective, Yongmin Liu and Peter R. Wiecha are invited to contribute two reviews on how to interface photonics with artificial intelligence, with focus on the inverse design strategy and applications beyond inverse design [6,7].

Deep learning could also help to deepen our understanding of complex nanophotonic structures. Emerging complex photonic structures derive their properties from a large network of inter-dependent nano-elements with both local and global connections. The vast parameter space offers unprecedented opportunities for device application, but at the same time presents a daunting challenge for developing an understanding of such complex structures. Deep learning could leverage such problems to recognize misaligned hyperfine orbital angular momentum modes [8], achieve phase compensation for free-space angular momentum-encoded quantum key distribution [9], enable automatic highly efficient photon distribution estimation via deep reinforcement learning [10], engineer multiple bound states in the continuum by latent representation of free-form structures [11], and identify modulation format in fiber communications using a single dynamical node [12].

The ever-increasing computing power required by deep learning prompts the search for alternative computing methods that are faster and more energy efficient. Optical analog computing can be passive with minimal energy consumption, and more importantly, its intrinsic parallelism can significantly accelerate computing speed. Such a capability is potentially useful as an optical preprocessor for real-time high-throughput image processing. Wave dynamics in a highly scattering medium can provide an alternative way to perform neuromorphic computing. Nanostructured photonic devices can exploit sub-wavelength linear and nonlinear scatterers to realize complex input-output mapping far beyond the capabilities of traditional nanophotonic devices. Theoretical and experimental exploration works in this feature issue have discussed neuromorphic node based on quantum dot laser [13], end-to-end optical

backpropagation for training neural networks [14], free-space optical neural network based on thermal atomic nonlinearity [15], deep-learning empowered dynamic wavefront shaping in nonstationary scattering media [16], delay weight plasticity based on supervised learning in photonic spiking neural networks [17], and all-optical neuromorphic binary convolution with spiking VCSEL neurons [18].

Another area of great interest is the application of deep learning in imaging analysis and computational cameras. Representative examples include deep compressive imaging techniques via optimized-pattern scanning [19], compressed ultrafast photography via an augmented-Lagrangian and deep-learning hybrid algorithm [20], and deep plug-and-play priors for spectral snapshot compressive imaging [21]. Image reconstruction techniques include denoising and reconstruction of super-resolution structured illumination microscopy images [22], and a simple low-SWaP hybrid machine vision system for universal training and generalized image reconstruction [23]. Novel imaging strategies through unknown scattering media based on physics informed learning [24], and incoherent imaging through highly dynamic and optically thick turbid media [25] are also demonstrated.

The last area that this feature issue highlights is the use of deep learning for sensors such as a smart ring resonator-based sensor for multicomponent chemical analysis [26] and sensing in the presence of strong noise by deep learning of dynamic multimode fiber interference [27]. A real-time deep learning design tool for far-field radiation profile [28] has been demonstrated for easy adoption and accessibility for device designers.

We take this opportunity to thank Prof. Lan Yang, the Editor-in-Chief, and Prof. Cun-Zheng Ning, the Deputy Editor for inviting us to organize this feature issue. They provide many helpful suggestions that greatly improve this feature issue. We also want to thank the editorial staff for their great support. We are also grateful to the reviewers that have helped us assess the quality and originality of the large amount of submitted papers, while keeping a high standard for our feature issue. Finally, and most importantly, we thank all the authors for submitting and contributing to this feature issue which provides a snapshot of exciting ongoing research work in the field and the opportunities that deep learning tools offer for the future of photonics technologies.

## REFERENCES

- S. So, Y. Yang, T. Lee, and J. Rho, "On-demand design of spectrally sensitive multiband absorbers using an artificial neural network," *Photon. Res.* **9**, B153–B158 (2021).
- P. Dai, Y. Wang, Y. Hu, C. H. de Groot, O. Muskens, H. Duan, and R. Huang, "Accurate inverse design of Fabry-Pérot-cavity-based color filters far beyond sRGB via a bidirectional artificial neural network," *Photon. Res.* **9**, B236–B246 (2021).
- Z. Zhen, C. Qian, Y. Jia, Z. Fan, R. Hao, T. Cai, B. Zheng, H. Chen, and E. Li, "Realizing transmitted metasurface cloak by a tandem neural network," *Photon. Res.* **9**, B229–B235 (2021).
- C. Liu, W. M. Yu, Q. Ma, L. Li, and T. J. Cui, "Intelligent coding metasurface holograms by physics-assisted unsupervised generative adversarial network," *Photon. Res.* **9**, B159–B167 (2021).
- Y. Ren, L. Zhang, W. Wang, X. Wang, Y. Lei, Y. Xue, X. Sun, and W. Zhang, "Genetic-algorithm-based deep neural networks for highly efficient photonic device design," *Photon. Res.* **9**, B247–B252 (2021).
- Y. Xu, X. Zhang, Y. Fu, and Y. Liu, "Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks," *Photon. Res.* **9**, B135–B152 (2021).
- P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photon. Res.* **9**, B182–B200 (2021).
- X. Wang, Y. Qian, J. J. Zhang, G. Ma, S. Zhao, R. F. Liu, H. Li, P. Zhang, H. Gao, F. Huang, and F. Li, "Learning to recognize misaligned hyperfine orbital angular momentum modes," *Photon. Res.* **9**, B81–B86 (2021).
- X. Wang, T. Wu, C. Dong, H. Zhu, Z. Zhu, and S. Zhao, "Integrating deep learning to achieve phase compensation for free-space orbital-angular-momentum-encoded quantum key distribution under atmospheric turbulence," *Photon. Res.* **9**, B9–B17 (2021).
- J. Ma, Z. Piao, S. Huang, X. Duan, G. Qin, L. Zhou, and Y. Xu, "Monte Carlo simulation fused with target distribution modeling via deep reinforcement learning for automatic high-efficiency photon distribution estimation," *Photon. Res.* **9**, B45–B56 (2021).
- R. Lin, Z. Alnakhli, and X. Li, "Engineering of multiple bound states in the continuum by latent representation of freeform structures," *Photon. Res.* **9**, B96–B103 (2021).
- Q. Cai, Y. Guo, P. Li, A. Bogris, K. A. Shore, Y. Zhang, and Y. Wang, "Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing," *Photon. Res.* **9**, B1–B8 (2021).
- G. Sarantoglou, M. Skontranis, A. Bogris, and C. Mesaritakis, "Experimental study of neuromorphic node based on a multiwaveband emitting two-section quantum dot laser," *Photon. Res.* **9**, B87–B95 (2021).
- X. Guo, T. D. Barrett, Z. M. Wang, and A. I. Lvovsky, "Backpropagation through nonlinear units for the all-optical training of neural networks," *Photon. Res.* **9**, B71–B80 (2021).
- A. Ryou, J. Whitehead, M. Zhelyeznyakov, P. Anderson, C. Keskin, M. Bajcsy, and A. Majumdar, "Free-space optical neural network based on thermal atomic nonlinearity," *Photon. Res.* **9**, B128–B134 (2021).
- Y. Luo, S. Yan, H. Li, P. Lai, and Y. Zheng, "Towards smart optical focusing: deep learning empowered dynamic wavefront shaping through nonstationary scattering media," *Photon. Res.* **9**, B262–B278 (2021).
- Y. Han, S. Xiang, Z. Ren, C. Fu, A. Wen, and Y. Hao, "Delay-weight plasticity-based supervised learning in optical spiking neural networks," *Photon. Res.* **9**, B119–B127 (2021).
- Y. Zhang, J. Robertson, S. Xiang, M. Hejda, J. Bueno, and A. Hurtado, "All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes," *Photon. Res.* **9**, B201–B209 (2021).
- K. Zhang, J. Hu, and W. Yang, "Deep compressed imaging via optimized pattern scanning," *Photon. Res.* **9**, B57–B70 (2021).
- C. Yang, Y. Yao, C. Jin, D. Qi, F. Cao, Y. He, J. Yao, P. Ding, L. Gao, T. G. Jia, J. Liang, Z. Sun, and S. Zhang, "High-fidelity image reconstruction for compressed ultrafast photography via an augmented-Lagrangian and deep-learning hybrid algorithm," *Photon. Res.* **9**, B30–B37 (2021).
- S. Zheng, Y. Liu, Z. Meng, M. Qiao, Z. Tong, X. Yang, S. Han, and X. Yuan, "Deep plug-and-play priors for spectral snapshot compressive imaging," *Photon. Res.* **9**, B18–B29 (2021).
- Z. H. Shah, M. Müller, T.-C. Wang, P. M. Scheidig, A. Schneider, M. Schüttelpelz, T. Huser, and W. Schenck, "Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images," *Photon. Res.* **9**, B168–B181 (2021).
- B. Muminov, A. Perry, R. Hyder, M. S. Asif, and L. T. Vuong, "Toward simple, generalizable neural networks with universal training for low-SWaP hybrid vision," *Photon. Res.* **9**, B253–B261 (2021).
- S. Zhu, E. Guo, J. Gu, L. Bai, and J. Han, "Imaging through unknown scattering media based on physics-informed learning," *Photon. Res.* **9**, B210–B219 (2021).

25. S. Zheng, H. Wang, S. Dong, F. Wang, and G. Situ, "Incoherent imaging through highly nonstatic and optically thick turbid media based on neural network," *Photon. Res.* **9**, B220–B228 (2021).
26. Z. Li, H. Zhang, B. T. T. Nguyen, S. Luo, P. Y. Liu, J. Zou, Y. Shi, H. Cai, Z. Yang, Y. Jin, Y. Hao, Y. Zhang, and A.-Q. Liu, "Smart ring resonator-based sensor for multicomponent chemical analysis via machine learning," *Photon. Res.* **9**, B38–B44 (2021).
27. L. V. Nguyen, C. C. Nguyen, G. Carneiro, H. Ebendorff-Heidepriem, and S. C. Warren-Smith, "Sensing in the presence of strong noise by deep learning of dynamic multimode fiber interference," *Photon. Res.* **9**, B109–B118 (2021).
28. J. Qie, E. Khoram, D. Liu, M. Zhou, and L. Gao, "Real-time deep learning design tool for far-field radiation profile," *Photon. Res.* **9**, B104–B108 (2021).

# PHOTONICS Research

## Modulation format identification in fiber communications using single dynamical node-based photonic reservoir computing

QIANG CAI,<sup>1,†</sup> YA GUO,<sup>1,2,†</sup> PU LI,<sup>1,3,4,\*</sup> ADONIS BOGRIS,<sup>5</sup> K. ALAN SHORE,<sup>6</sup>  
YAMEI ZHANG,<sup>7</sup> AND YUNCAI WANG<sup>3</sup>

<sup>1</sup>Key Laboratory of Advanced Transducers and Intelligent Control System, Ministry of Education, Taiyuan University of Technology, Taiyuan 030024, China

<sup>2</sup>School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710072, China

<sup>3</sup>School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China

<sup>4</sup>Key Laboratory of Specialty Fiber Optics and Optical Access Networks, Shanghai University, Shanghai 200444, China

<sup>5</sup>Department of Informatics and Computer Engineering, University of West Attica, Athens 12243, Greece

<sup>6</sup>School of Electronic Engineering, Bangor University, Wales LL57 1UT, UK

<sup>7</sup>Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

\*Corresponding author: lipu8603@126.com

Received 9 September 2020; revised 12 November 2020; accepted 13 November 2020; posted 18 November 2020 (Doc. ID 409114); published 24 December 2020

We present a simple approach based on photonic reservoir computing (P-RC) for modulation format identification (MFI) in optical fiber communications. Here an optically injected semiconductor laser with self-delay feedback is trained with the representative features from the asynchronous amplitude histograms of modulation signals. Numerical simulations are conducted for three widely used modulation formats (on-off keying, differential phase-shift keying, and quadrature amplitude modulation) for various transmission situations where the optical signal-to-noise ratio varies from 12 to 26 dB, the chromatic dispersion varies from -500 to 500 ps/nm, and the differential group delay varies from 0 to 20 ps. Under these situations, final simulation results demonstrate that this technique can efficiently identify all those modulation formats with an accuracy of >95% after optimizing the control parameters of the P-RC layer such as the injection strength, feedback strength, bias current, and frequency detuning. The proposed technique utilizes very simple devices and thus offers a resource-efficient alternative approach to MFI. © 2020 Chinese Laser Press

<https://doi.org/10.1364/PRJ.409114>

### 1. INTRODUCTION

Fiber-optic communication systems are expected to be capable of adaptively adjusting various transmission parameters such as modulation formats, line rates, and spectrum assignments, based on the varying channel conditions and traffic demands in order to maximize the spectral and energy efficiencies [1–4]. The dynamic variation of transmission parameters imposes new requirements for the optical receivers in such elastic optical networks (EONs). To demodulate the transmission signal at the digital receivers, one must know the type of modulation format. Consequently, correct identification of modulation formats is rather crucial for high-quality communication [5,6].

The feature-based (FB) approach is an effective way to achieve modulation format identification (MFI) [7–11], carried out by using different tools to analyze the associated feature parameters from transmission signals. For instance, Nandi *et al.*

completed an identification of amplitude modulation (AM), frequency modulation (FM), M-ary amplitude shift-keying (MASK), and M-ary frequency-shift keying (MFSK) signals by analyzing their instantaneous phase and frequency information using the decision tree algorithm [9]. Park *et al.* realized the identification of MASK, MFSK, and M-ary phase-shift keying (MPSK) signals by employing the support vector machine to analyze the frequency features of modulated signals [10]. Khan *et al.* confirmed that using artificial neural network (ANN) can achieve the identification of six widely used modulation formats [including on-off keying (OOK), differential phase-shift keying (DPSK), M-ary quadrature amplitude modulation (MQAM), etc.] through analyzing their amplitude features [11].

Among them, ANN-based MFI technologies especially attract great attention due to their enormous calculation power and high accuracy. Typically, Wong *et al.* identified commonly

used MASK, MFSK, and MQAM signals using a multilayer perceptron (MLP) neural network [12]. O’Shea *et al.* successfully realized the identification of OOK, BPSK, MASK, MPSK, and MQAM signals by analyzing higher-order statistics with a convolutional neural network (CNN) [13]. Wang *et al.* implemented MFI on quadrature phase-shift keying (QPSK), phase-shift keying (PSK), and MQAM signals through a constellation diagram employing a deep neural network (DNN) with multiple nonlinear layers [14]. However, it should be pointed that a typical ANN architecture consists of at least three layers (i.e., the input layer, the hidden layer, and the output layer) with neuron nodes between two adjacent layers interlinked by variable trained weights.

Time-delayed reservoir computing (RC) [15] is a new type of ANN consisting of one nonlinear node under delayed feedback. For time-delayed RC, the nonlinear component under delayed feedback is used as the reservoir. Moreover, the output layer weight in the whole system is the only part that needs to be trained, and the process can be completed with low complexity.

Up to now, there have been many RCs reported based on optoelectronic or all-optical devices [16–30]. As far as we are aware of, the RC systems can complete complex computational tasks with high performance such as spoken digit recognition [17], nonlinear time series prediction [18], and wireless channel equalization [19]. In particular, photonic reservoir computing (P-RC) based on semiconductor lasers with time-delayed feedback is very promising for high-speed implementation of the RC. For instance, the P-RC performs the identification and classification of a packet header for switching in an optical network application [31,32]. In addition, the P-RC has been also proposed to address signal recovery in optical communication systems [33,34].

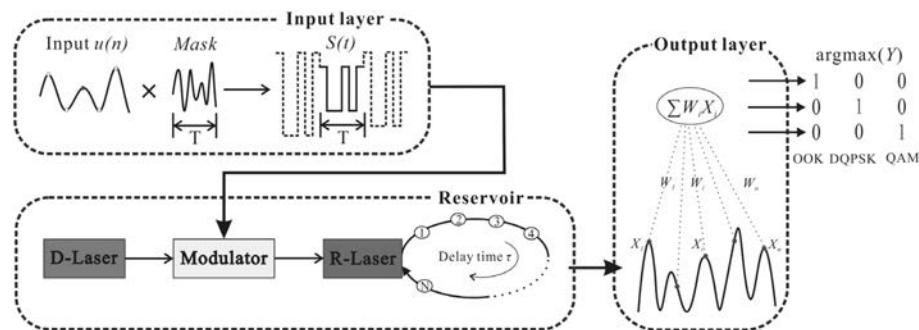
In this paper, the P-RC with semiconductor lasers is, for the first time to our knowledge, introduced to the field of MFI. After extracting representative features of amplitude histograms through asynchronous sampling OOK, differential quadrature phase-shift keying (DQPSK), and 16 quadrature amplitude modulation (16QAM) signals, we numerically implement correct identification results by means of the P-RC constructed with a delay-feedback semiconductor laser. Herein, we point

out that memory is a typical advantage of RC, especially compared to a feedforward NN. Memory in an RC originates from the recurrent network connections, allowing information to remain in the network over finite time. Past information therefore mixes with the current input. In our photonic RC, optical delayed feedback introduces recurrences resulting in the ring topology, so that the optical reservoir possesses this memory property. As the input of the reservoir in our task, the envelope of the asynchronous amplitude histogram from any modulation signal is almost continuous and shows a short-time relevance. Therefore, the memory in our RC needs to vanish after some time to allow responses to be influenced only by the recent past. This property is referred to as fading memory [35] and plays an important role in improving the identification performance.

Experiences demonstrate that the reservoir layer in the absence of input information should work in the nonlinear region to exhibit sufficiently different dynamical responses to input with different classes [20]. To accurately fix the corresponding nonlinear regime, we examine these critical hyperparameters (i.e., injection strength, feedback strength, response laser current, and frequency detuning) in a reasonable range based on the bifurcation diagram of the semiconductor lasers. Besides, based on our research on the nonlinear dynamics of semiconductor lasers [36,37], we find that the order of the hyperparameters does not matter. After that, we finally achieve an MFI with an overall estimation accuracy as high as 95%.

## 2. THEORETICAL MODEL

Figure 1 shows the schematic of the MFI based on the P-RC with semiconductor lasers. The whole system consists of three parts: input layer, reservoir, and output layer. In our implementation, the various modulation formats (OOK, DQPSK, QAM) commonly used in long-haul optical communication systems obtained their amplitude feature  $u(n)$  through asynchronous sampling. Specifically, the input  $u[n]$  is generated in the following way: 10 Gbps OOK, 40 Gbps DQPSK, and 100 Gbps 16QAM signals transmit over the emulated communication channel with different optical signal-to-noise



**Fig. 1.** Schematic of the MFI based on the P-RC with semiconductor lasers. This system consists of three parts: input layer, reservoir, and output layer. The input  $u(n)$  is multiplied by a mask with a period of  $T$ , and then the resulting stream  $S(t) = \text{Mask} \times u(n)$  is fed into the reservoir through a modulator. The reservoir is a master-slave configuration constructed by a response laser (R-Laser) with a self-delay feedback loop injected by a drive laser (D-Laser). Note that there are  $N$  virtual nodes at each interval  $\theta$  in the feedback loop with a delay time of  $T$ . The transient states of the R-Laser  $X_i$  are read out for training the connection weights  $W_i$  between the reservoir and the output layer. The final output nodes are weighted by the sums of the transient states  $\sum X_i W_i$ .

ratios (OSNRs)/chromatic dispersions (CDs)/differential group delays (DGDs) and are then asynchronously sampled at a sampling rate of 500 MSa/s in a duration time of 0.2 ms to obtain 100,000 amplitude samples. After normalizing all of the sampled amplitudes, we form the associated asynchronous amplitude histograms (AAHs) containing 100 bins as shown in the following Fig. 3. Finally, 11,700 sample data sets are generated and used as the input information  $u[n]$ . The input information is a one-dimensional data vector  $u(n)$ , where  $n \in \mathbb{Z}$  is the discrete time. In the input layer,  $u[n]$  is preprocessed and multiplied by a mask sequence, which plays the role of random weight connections from the input to the reservoir layer. Note that different kinds of mask sequences can induce different effects on the RC performance. In the simulation, we actually investigated three kinds of commonly used mask signals, which are binary [38], six-level [39], and chaos masks [40], respectively. After analyzing their effect on the identification performance in the following Section 3, we finally select the chaos signal as the mask for RC. The amplitude feature  $u(n)$  first undergoes a sample-hold operation, where each sampling point has a duration of  $T$  and then is multiplied by a chaos mask sequence with a length of  $T$ . The resulting sequence  $S(t) = \text{Mask} \times u(n)$  is further injected into the following photonic reservoir.

In the reservoir, the optical signal from the drive laser (D-Laser) is modulated by the loaded signal  $S(t)$  from the input layer and then injected into the response laser (R-Laser) with a time-delayed optical feedback. Following ideas introduced in RC with the delay system, the input information generates nonlinear transient states in the context of previous input responses. This is because the induced transient states at time  $t$  depend on the output of the nonlinear node within the time interval  $[t - \tau, t]$ , with  $\tau$  being the delay time. In addition, the dynamics of the delay system exhibit the properties of high dimensionality and short-term memory, which accord well with the requirements of the P-RC. Within one delay interval of length  $\tau$ , the feedback delay loop contains  $N$  virtual nodes at each interval time  $\theta$  ( $\theta = \tau/N$ ). We denote these  $N$  equidistant points as “virtual nodes,” as their roles are actually analogous to the nodes of a traditional reservoir. The values of the delayed variable at each of the  $N$  points define the states of the virtual nodes, which show the transient response of the reservoir when the input information is at the specific time. The whole working process in the reservoir can be modeled by Eqs. (1) and (2) [30,40,41]:

$$\begin{aligned} \frac{dE(t)}{dt} &= \frac{1 + i\alpha}{2} \left\{ \frac{g[N(t) - N_0]}{1 + \epsilon|E(t)|^2} - \frac{1}{\tau_p} \right\} E(t) \\ &+ \frac{k_f}{\tau_{\text{in}}} E(t - \tau) \exp(-i2\pi\nu\tau) \\ &+ \frac{k_{\text{inj}}}{\tau_{\text{in}}} E_{\text{inj}}(t) \exp(i2\pi\Delta\nu t) + \sqrt{2\beta N(t)} \chi(t), \quad (1) \end{aligned}$$

$$\frac{dN(t)}{dt} = J - \frac{N(t)}{\tau_s} - \frac{g[N(t) - N_0]}{1 + \epsilon|E(t)|^2} |E(t)|^2, \quad (2)$$

where  $E$  is the slowly varying complex electric field and  $N$  is the average carrier density. The parameters  $k_f$  and  $k_{\text{inj}}$  represent the

feedback strength of external cavity of the R-Laser and the injection strength from the D-Laser to the R-Laser, respectively.  $\nu$  denotes the frequency of the free-running R-Laser, and  $\Delta\nu$  denotes the frequency detuning from the D-laser to the R-laser.  $J$  is the injection current. The carrier density at transparency  $N_0 = 4.55 \times 10^{23} \text{ m}^{-3}$ , the linewidth-enhancement factor  $\alpha = 5.0$ , the differential gain coefficient  $g = 1.414 \times 10^{-12} \text{ m}^3 \cdot \text{s}^{-1}$ , the gain saturation coefficient  $\epsilon = 2.0 \times 10^{-23}$ , the internal cavity round-trip time  $\tau_{\text{in}} = 7.38 \text{ ps}$ , the photon lifetime  $\tau_p = 1.17 \text{ ps}$ , and the carrier lifetime  $\tau_s = 2.5 \text{ ns}$ , respectively. The feedback delay time  $\tau$  is 8 ns, and duration  $T = \tau \cdot \chi(t)$  is a white Gaussian noise with zero mean and unity variance, used to model the spontaneous emission noise.  $\beta$  is the strength of the spontaneous emission noise. [Note that the following simulation results are obtained at  $\beta = 0$ , except for the results in Table 3, where  $\beta = 1.5 \times 10^{-6}$ .] Considering that the input signal is multiplied with the mask being used to modulate the optical signal through phase modulation, the injected slowly varying complex electric field  $E_{\text{inj}}$  is written as Eq. (3):

$$E_{\text{inj}}(t) = \sqrt{I_d} \exp[i\pi S(t)], \quad (3)$$

where  $I_d$  is the photon number of continuous-wave output from the D-Laser,  $I_d = 3.757 \times 10^{20}$ .  $S(t)$  represents the masked input signal.

In the output layer, the transient states of the R-Laser  $X_i$  are read out for training the connection weights  $W_i$  between the reservoir and the output layer as defined in Eq. (4). Specifically, the optimal readout weights are calculated using the ridge regression algorithm in our work. The final output target  $Y(n)$  of three identified modulation formats is provided by using one-hot encoding [42] as follows. In the training stage of the P-RC, each input vector  $u(n)$  has a corresponding  $M \times 1$  binary vector  $Y$  with only one nonzero element. The location of ‘1’ in  $Y$  indicates the signal modulation format type:

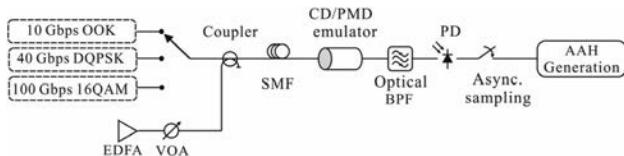
$$Y(n) = \sum X_i W_i. \quad (4)$$

Once the training process of the P-RC is completed, we employ the error rate (ER) to evaluate the MFI performance by a separate set of data called testing data set as defined in Eq. (5). In the testing stage, the location of the largest element in each corresponding output vector  $v$ ,  $\text{argmax}\{v\}$ , is used as an identifier of the signal modulation format. The identified modulation formats are compared with true ones, which are provided by labels  $Y$  of the testing data set. Herein,  $Q$  represents the total of the samples of the testing data set, while  $b$  is the number of erroneous identification samples:

$$\text{ER} = \frac{b}{Q} \times 100\%. \quad (5)$$

### 3. NUMERICAL RESULTS AND DISCUSSIONS

As mentioned before, we choose 10 Gbps OOK, 40 Gbps DQPSK, and 100 Gbps 16QAM modulation format signals commonly used in long-haul fiber communication systems to validate the feasibility of our MFI proposal based on the P-RC in simulations. As shown in Fig. 2, these signals with three modulation formats transmit over a 30 km standard single-mode

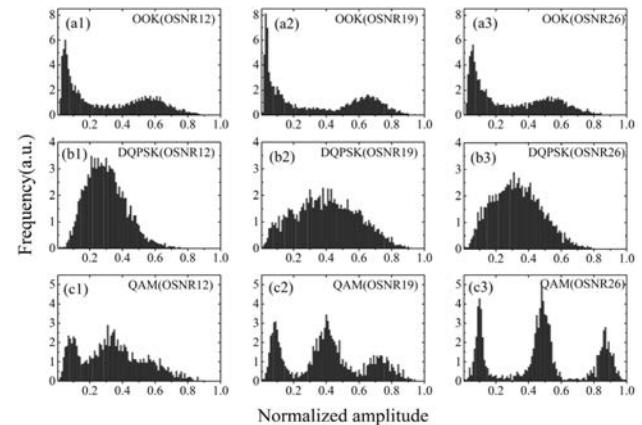


**Fig. 2.** Sketch of the emulated transmission system for asynchronous amplitude histogram generation. EDFA, erbium-doped fiber amplifier; SMF, single-mode fiber; CD/PMD, chromatic dispersion/polarization mode dispersion; BPF, band-pass filter; PD, photodetector; AAH, asynchronous amplitude histogram.

fiber (SMF; dispersion coefficient is  $16 \text{ ps}/(\text{nm} \cdot \text{km})$ , attenuation factor is  $0.2 \text{ dB/km}$ , and nonlinear coefficient is  $1.27 \text{ km}^{-1} \cdot \text{W}^{-1}$ ), respectively. An erbium-doped fiber amplifier (EDFA) is used to add amplified spontaneous emission (ASE) noise into the signals, while a variable optical attenuator (VOA) is used to adjust the OSNRs. A CD and a polarization mode dispersion (PMD) emulator are used to introduce variable amounts of CDs and DGDs into the signal, respectively. After removing the redundant noise by an optical band-pass filter (BPF) and direct detection by a photodetector (PD), the resulting electrical signal is asynchronously sampled to form asynchronous amplitude histograms (AAHs). The OSNR ranges considered in this work are the ones used in practice for reliable data transmission with the abovementioned signal types. There are various losses in the signal transmission process of optical fiber. Specifically, we investigate the three types of signals in the following scenarios: (i) their OSNRs are adjusted, ranging from 12 to 26 dB through a VOA at a step of 1 dB; (ii) their CDs are set, ranging from  $-500$  to  $500 \text{ ps/nm}$  via a CD emulator at a step of  $80 \text{ ps/nm}$ ; (iii) their DGDs are located by a PMD emulator in the range of 0–20 ps with a step of 5 ps. Here we want to point out that using a fixed length of standard SMF with different OSNR and CD following is a common method to simulate the transmission system for asynchronous amplitude histogram generation [11,43]. In this way, the OSNR and CD can be easily adjusted to obtain associated data. Thus, we also apply this method in our simulation. Note that the nonlinear effects are very weak in the 30 km fiber propagation, and thus their impact is almost negligible.

Based on the above scenarios, 11,700 samples of modulation format signals in total are collected, and the randomly selected subsets of this large data set are then used for training and testing the performance of the P-RC. The examples of amplitude histograms for OOK, DQPSK, and 16QAM signals after asynchronous sampling are typically shown in Fig. 3. Herein, the OSNRs in each column are 12, 19, and 26 dB from left to right, but the corresponding CD and the DGD are fixed at  $80 \text{ ps/nm}$  and 5 ps, respectively. From each column, it can be clearly seen that different modulation formats exhibit distinct shapes. Even though the associated shapes change with various OSNRs from the first column to the third column, three types of modulation formats maintain distinct characteristics for each format. Therefore, the characteristic features of the asynchronous amplitude histograms have been widely accepted to be exploited for MFI.

In the following, we will use the features from the asynchronous amplitude histograms to train and test the P-RC so that

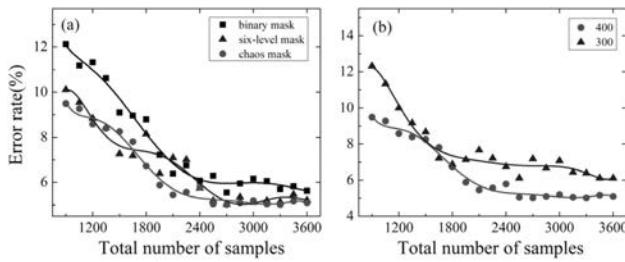


**Fig. 3.** Typical asynchronous amplitude histograms for (a1)–(a3) OOK, (b1)–(b3) DQPSK, and (c1)–(c3) QAM formats after propagation through the emulated communication channel. From left to right, each column has an OSNR of 12, 19, and 26 dB, while the corresponding CD and the DGD are fixed at  $80 \text{ ps/nm}$  and 5 ps, respectively.

MFI can be realized. At the first stage, we confirm the appropriate number of samples and the optimized size of the virtual nodes in the P-RC for high efficiency. In our scheme, we adopt the commonly used  $k$ -fold cross-validation algorithm [44] to eliminate the impact of the specific division of the available data samples between training and testing. This means that the entire process of training and testing is repeated  $k$  times (with  $k = 5$ ) on the same data, but each time with a different assignment of data samples to each of the two stages. The final ER in testing on the  $y$  axis is the mean across these  $k = 5$  runs. We compared the identification results with different types of mask signals (the binary mask, six-level mask, and chaos mask) by varying the sample numbers. The binary mask consists of a piecewise constant function with a randomly modulated binary sequence  $\{-1, 1\}$ . Similar to the binary mask, the six-level mask is composed of a random sequence  $\{\pm 1, \pm 0.6, \pm 0.3\}$ . The chaos mask signal is generated from another semiconductor laser with optical feedback [45]. Here the amplitude of the chaos mask is rescaled so that the standard deviation of the chaos mask is set to 1 and the mean value is set to 0. As shown in Fig. 4(a), the ER tends to be a stationary value when the sample number reaches around 2700. Meanwhile, it is clear that the chaos mask performs better than the other two types of masks. Thus, we select the chaos signal as the final mask in the following simulation. On the other hand, it can be observed that no matter whether the number of virtual nodes  $N$  in the P-RC is set as 300 or 400, the ER exhibits the same changing trend. Meanwhile, it should be noted that the ER corresponding to the virtual node number  $N = 400$  is lower than that in the other cases of  $N = 300$ .

As shown in Fig. 4(b), the ER decreases with the increase of the sample number and tends to be a stationary value. The red curve is the corresponding red curve (chaos mask) in Fig. 4(a). Consequently, 2700 samples of modulation signals and the virtual node number  $N = 400$  are selected in our simulation.

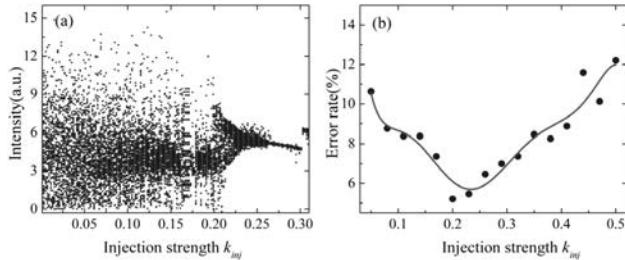
Furthermore, we analyze the effect of the four key parameters in the P-RC layer on the MFI performance, which are the



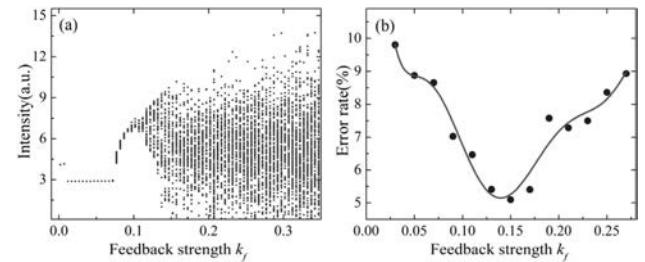
**Fig. 4.** (a) Identification error rate on the total (training and test) sample numbers of the binary mask (black), the six-level mask (blue), and the chaos mask signals (red). (b) Dependence of the identification error rate on the total (training and test) sample numbers at different virtual node sizes of 300 (blue) and 400 (red).

injection strength  $k_{\text{inj}}$  from the D-laser to R-laser, feedback strength  $k_f$  of the R-laser, bias current  $I_R$  of the R-laser, and frequency detuning  $\Delta\nu$  from the D-laser to R-laser. Figure 5 depicts the associated influence of the injection strength. When there is no input, the P-RC is a typical master-slave laser configuration. From its bifurcation diagram [Fig. 5(a)], one can observe that when the injection strength  $k_{\text{inj}}$  is less than 0.21, the output of the P-RC layer is in a chaotic state. With increase of the injection strength  $k_{\text{inj}}$ , the P-RC laser changes from the periodic oscillation state to a single-cycle state. On the other hand, Fig. 5(b) shows the dependence of the identification ER on the injection strength when the input is considered. It can be seen from Fig. 5(b) that the ER is first decreased but then increased with the increase of the injection strength  $k_{\text{inj}}$ . The lowest ER of 5.21% can be obtained when the optimum value of injection strength  $k_{\text{inj}}$  is equal to 0.2. Comparing its corresponding bifurcation diagram [Fig. 4(a)], it can be found that when  $k_{\text{inj}}$  is around 0.2, the delay-coupled semiconductor laser reservoir system in the absence of input works at the edge of the chaos region. This phenomenon is consistent with that in Ref. [15], where electrical reservoir computing based on a single dynamical node is used for standard benchmarking tasks such as spoken digit recognition and nonlinear time series prediction.

Figure 6 shows the effect of the feedback strength  $k_f$  of the R-laser on the identification performance. Here the injection strength  $k_{\text{inj}}$  is set at 0.2, while other parameters remain unchanged ( $I_R = 1.3I_{\text{th}}$  and  $\Delta\nu = -10$  GHz). When the feed-



**Fig. 5.** (a) Bifurcation diagram of the output optical intensity versus the injection strength  $k_{\text{inj}}$  for  $k_f = 0.18$ ,  $I_R = 1.3I_{\text{th}}$ , and  $\Delta\nu = -10$  GHz. (b) Identification error rate (ER) at different injection strengths  $k_{\text{inj}}$  (blue dots), while the red curve is plotted by executing a sliding window averaging to the associated data points.

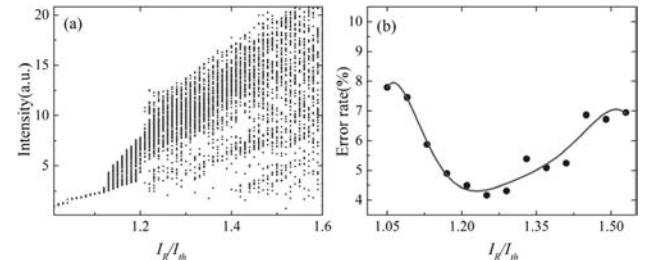


**Fig. 6.** (a) Bifurcation diagram of the output optical intensity versus the feedback strength  $k_f$  for  $k_{\text{inj}} = 0.2$ ,  $I_R = 1.3I_{\text{th}}$ , and  $\Delta\nu = -10$  GHz. (b) Identification ER at different feedback strengths  $k_f$  (blue dots), while the red curve is plotted by executing a sliding window averaging to the associated data points.

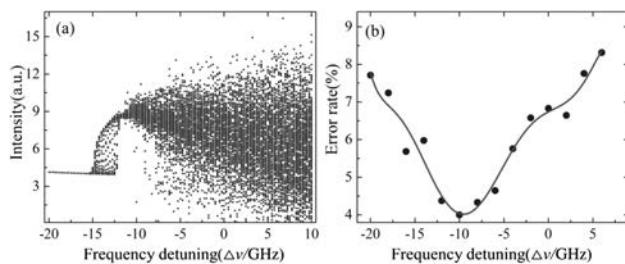
back strength  $k_f$  varies from 0.01 to 0.35, it can be seen from the corresponding bifurcation diagram in Fig. 6(a) that the output of the reservoir without the input follows a quasiperiodic route to chaos with the increase of the feedback strength  $k_f$ . After the input is considered, the final identification ER of our P-RC in Fig. 6(b) is decreased at first and then increased with increasing feedback strength  $k_f$ . When the feedback strength  $k_f$  is 0.15, the ER reaches a minimum value of 5.09%. When the feedback strength grows, the system is in the chaotic state and is more sensitive to changes in initial conditions. So the identification performance will decline.

Figure 7 illustrates the influence of the bias current  $I_R$  of the R-laser in the system. In the analysis, the injection strength  $k_{\text{inj}}$  and the feedback strength  $k_f$  are set to be 0.2 and 0.15, respectively. Figure 7(a) is the corresponding bifurcation diagram of the output optical intensity versus the bias current  $I_R$  of R-Laser. As can be seen, the RC subsystem will gradually leave a single-cycle state to the chaotic region when the bias current of the R-Laser increases from  $1.01I_{\text{th}}$  to  $1.6I_{\text{th}}$ . Note that  $I_{\text{th}}$  corresponds to the threshold current of the R-Laser. On the other hand, the identification ER in Fig. 7(b) first decreases and then increases with the increase of  $I_R$ . Finally, a minimum ER of 4.25% can be obtained when the bias current of the R-Laser equals  $1.25I_{\text{th}}$ .

Figure 8 investigates the influence of the frequency detuning  $\Delta\nu$  between the D-Laser and the R-Laser on the identification results. Figure 8(a) gives the bifurcation diagram as a function of the frequency detuning  $\Delta\nu$ . In the bifurcation diagram, the



**Fig. 7.** (a) Bifurcation diagram of the output optical intensity versus the bias current of the R-Laser  $I_R$  for  $k_{\text{inj}} = 0.2$ ,  $k_f = 0.15$ , and  $\Delta\nu = -10$  GHz. (b) Identification ER at different bias currents of the R-Laser  $I_R$  (blue dots), while the red curve is plotted by executing a sliding window averaging to the associated data points.



**Fig. 8.** (a) Bifurcation diagram of the output optical intensity versus the frequency detuning  $\Delta\nu$  between the D-Laser and the R-Laser for  $k_{\text{inj}} = 0.2$ ,  $k_f = 0.15$ , and  $I_R = 1.25I_{\text{th}}$ . (b) Identification ER at different frequency detunings  $\Delta\nu$  (blue dots), while the red curve is plotted by executing a sliding window averaging to the associated data points.

RC without the input can be observed in a single-cycle state when the frequency detuning  $\Delta\nu$  is  $-15$  GHz. With the increase of the frequency detuning  $\Delta\nu$ , it starts to enter a chaotic state at  $\Delta\nu = -10$  GHz. Figure 8(b) shows that the identification ER exhibits the same changing trend as Figs. 5(b), 6(b), and 7(b). That is, the ER decreases first and then increases with the increase of frequency detuning  $\Delta\nu$ . When the frequency detuning  $\Delta\nu$  is  $-10$  GHz, the ER can be reduced to a lowest value of 4.07%. Obviously, this best identification result is obtained at the edge of the chaos region by contrast with the bifurcation diagram. In this state, the RC has an infinite dimensional space so the input signal is mapped to the higher dimensional space and achieves the optimal identification results.

After the above processes, all key parameters in our P-RC can be adjusted to reach the optimum state as follows. (i) The number of virtual nodes  $N$  is chosen to be 400. (ii) The injection strength  $k_{\text{inj}}$  from the D-laser to the R-laser is set to be 0.2. (iii) The feedback strength  $k_f$  of the R-laser is set as 0.15. (iv) The bias current  $I_R$  of the R-Laser is  $1.25I_{\text{th}}$ . (v) The frequency detuning  $\Delta\nu$  between the D-Laser and R-Laser is set to  $-10$  GHz.

Under these conditions, we utilize the P-RC system to identify 1000 test cases with each type of modulation format for a considerable range of OSNRs (12–26 dB), CDs ( $-500$ – $500$  ps/nm), and DGDs (0–20 ps). Table 1 lists the associated test results of the proposed MFI technique through combining the ridge regression algorithm with the laser-based P-RC system. From it, one can confirm that the identification accuracies of OOK, DQPSK, and 16QAM sig-

nals can reach 95.1%, 95.7%, and 95.5%, respectively. These simulation results demonstrate that the proposed technique can effectively classify the modulation formats in fiber communication with an overall estimation accuracy of more than 95% and also in the presence of various link impairments such as OSNR, CD, and DGD. Meanwhile, in order to check the performance of our proposal, we also analyzed the MFI accuracy using only the ridge regression algorithm. In this case, the whole system only contains two layers: the input and output layers. The whole training process is as follows. The resulting sequence  $S$  from the input layer and the target matrix  $Y$  are directly used to calculate the final output connection weight  $W$  by means of the ridge regression algorithm. That is, there is no nonlinear mapping induced by the reservoir layer. The associated results on identification accuracies for different modulation formats are shown in Table 2. It is evident that the identification accuracies of all three modulation formats in the system only using the ridge regression algorithm are less than 76%. Therefore, we confirm that the identification accuracies can be significantly improved based on the proposed P-RC MFI system in our work.

Further, we investigate the influence of spontaneous emission noise on the performance of the P-RC. Table 3 shows the identification accuracies for different modulation formats using the MFI technique through the P-RC system under a typical spontaneous emission  $\beta = 1.5 \times 10^{-6}$  [46]. Comparing Table 1 and Table 3, one can find that our P-RC is extremely robust against spontaneous emission noise. That can be explained by the fact that the modulation format identification is a classification task that only requires a winner-takes-all decision. Note that the noise from the photodetector is additive, which can be eliminated using balance detection technology. Thus, the detection noise is not considered in this work.

**Table 2. Identification Accuracies for Different Modulation Formats Using Only the Ridge Regression Algorithm (Without the Reservoir Layer in the System)<sup>a</sup>**

	Identified Modulation Formats			
	OOK	DQPSK	QAM	
Actual Modulation Formats	OOK DQPSK QAM	75.2% 15.7% 9.1%	8.2% 72.6% 19.2%	8.5% 16.7% 74.8%

<sup>a</sup>The overall MFI accuracy is less than 76%.

**Table 3. Identification Accuracies for Different Modulation Formats Using the MFI Technique Through the P-RC System with a Typical Noise Value of  $\beta = 1.5 \times 10^{-6}$ <sup>a</sup>**

	Identified Modulation Formats			
	OOK	DQPSK	QAM	
Actual Modulation Formats	OOK DQPSK QAM	95.0% 3.5% 1.5%	1.2% 95.5% 3.3%	1.8% 2.9% 95.3%

<sup>a</sup>The overall MFI accuracy is more than 95%.

**Table 1. Identification Accuracies for Different Modulation Formats Using the MFI Technique Through Our Laser-Based P-RC System<sup>a</sup>**

	Identified Modulation Formats			
	OOK	DQPSK	QAM	
Actual Modulation Formats	OOK DQPSK QAM	95.1% 3.2% 1.7%	1.4% 95.7% 2.9%	1.7% 2.8% 95.5%

<sup>a</sup>The overall MFI accuracy is more than 95%.

Finally, we want to point out that direct comparison with other works is difficult in modulation format identification as mentioned in Refs. [12,47]. There are two main reasons for this difficulty. (i) There is no single unified database available. When different modulation formats with different symbol rates are applied, different accuracies will result for NN-based modulation format identification. (ii) Identification accuracy is frequently reported as a function of OSNR and CD. There are no benchmarking systems universally agreed upon these indexes. For instance, Guesmi *et al.* reported an identification accuracy of 90% for 16 Gbps QPSK and 16QAM when the OSNR is 14–26 dB and the CD is 800 ps/nm [48]. Saif *et al.* showed an identification accuracy of 98% for 10 Gbps QPSK and 16QAM when OSNR is about 20 dB; however, when the symbol rate is enhanced into 20 Gbps, the associated identification accuracy is decreased to 90% [49]. Xiang *et al.* reported an identification accuracy of 100% for 28 Gbps QPSK and 16QAM when the CD is -60–60 ps/nm; but when the CD is in the other range, the identification accuracy is decreased below 90% [43].

There are at least two benefits to the P-RC. (i) High speed. Our P-RC is based on the all-optical nonlinearity of laser diodes. Such a system might even reach processing speeds at the level of 100 GHz using off-the-shelf photonic components for telecommunication applications [50]. (ii) Low energy consumption. At this stage we can only provide a conservative estimate, according to Ref. [16]. Our all-optical transient computing scheme, including all-optical data input and read-out hardware, would be of the order of 10 mJ per signal, compared with 2 J per signal required by a standard desktop computer.

In addition, the current accuracy level in this work is mainly constricted by the number of used nodes due to the limited sample size. Once the sample size is sufficiently large (that means more nodes are fully trained), there are two approaches to further achieve higher accuracy levels with P-RC. One is to improve the complexity of the mask [40]. In P-RC, the mask is equivalent to the input connection weight in the traditional RC. It can bring rich nonlinear states for P-RC and thus map the input data into a high-dimensional space. So it is expected that using a more complicated random signal (such as amplified spontaneous noise) as the mask may further enhance the identification accuracy. The other is to introduce the multiple feedback loops [30]. Usually, the P-RC system with multiple optical feedback results in a much more complex response of the R-laser compared to that with a single optical feedback. Consequently, higher dimensional transformation of the input signal and thus higher accuracy levels may be achieved with P-RC.

#### 4. CONCLUSION

In this paper, we propose a simple MFI technique for fiber-optic communications by using P-RC trained with amplitude features extracted from the OOK, DQPSK, and 16QAM signals through asynchronous sampling. Numerical results demonstrate over 95% identification accuracy for three widely used modulation formats in various situations under OSNRs (12–26 dB), CDs (-500–500 ps/nm), and DGDs

(0–20 ps). Considering the good performance and its conceptual simplicity, it can be expected that P-RC may provide an efficient tool for realizing the wide application of MFI in future fiber communication systems.

**Funding.** National Natural Science Foundation of China (61775158, 61961136002, 61927811, U19A2076, 61705159, 61805168, 17174343, 11904157); Program for Guangdong Introducing Innovative and Entrepreneurial Teams; Program for the Top Young Academic Leaders of High Learning Institutions of Shanxi; National Cryptography Development Fund (MMJJ20170127); China Postdoctoral Science Foundation (2018M630283, 2019T120197); Natural Science Foundation of Shanxi Province (201901D211116); STCSM (SKLSFO2018-03); Project of Key Laboratory of Radar Imaging and Microwave Photonics (Nanjing University of Aeronautics and Astronautics), Ministry of Education (RIMP2019002).

**Disclosures.** The authors declare no conflicts of interest.

<sup>†</sup>These authors contributed equally to this work.

#### REFERENCES

1. A. Nag, M. Tornatore, and B. Mukherjee, "Optical network design with mixed line rates and multiple modulation formats," *J. Lightwave Technol.* **28**, 466–475 (2010).
2. S. J. Yoo, A. Lord, M. Jinno, and O. Gerstel, "Elastic optical networking: a new dawn for the optical layer?" *IEEE Commun. Mag.* **50**, s12–s20 (2012).
3. T. Zhang, J. Wang, Q. Liu, J. Z. Zhou, J. Dai, X. Han, Y. Zhou, and K. Xu, "Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks," *Photon. Res.* **7**, 368–380 (2019).
4. Y. Q. Chang, H. Wu, C. Zhao, L. Shen, S. N. Fu, and M. Tang, "Distributed Brillouin frequency shift extraction via a convolutional neural network," *Photon. Res.* **8**, 690–697 (2020).
5. Z. Pan, C. Yu, and A. E. Willner, "Optical performance monitoring for the next generation optical communication networks," *Opt. Fiber Technol.* **16**, 20–45 (2010).
6. R. Borkowski, D. Zibar, and I. T. Monroy, "Anatomy of a digital coherent receiver," *IEICE Trans. Commun.* **E97-B**, 1528–1536 (2014).
7. E. E. Azzouz and A. K. Nandi, *Automatic Modulation Recognition of Communication Signals* (Kluwer Academic, 1996).
8. O. A. Dobre, A. Abdi, Y. Bar-Ness, and W. Su, "A survey of automatic modulation classification techniques: classical approaches and new trends," *IET Commun.* **1**, 137–156 (2007).
9. A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Trans. Commun.* **46**, 431–436 (1998).
10. C. S. Park, J. H. Choi, S. P. Nah, W. Jang, and D. Y. Kim, "Automatic modulation recognition of digital signals using wavelet features and SVM," in *10th International Conference on Advanced Communication Technology* (2008), pp. 387–390.
11. F. N. Khan, Y. Zhou, A. P. T. Lau, and C. Lu, "Modulation format identification in heterogeneous fiber-optic networks using artificial neural networks," *Opt. Express* **20**, 12422–12431 (2012).
12. M. L. D. Wong and A. K. Nandi, "Automatic digital modulation recognition using artificial neural network and genetic algorithm," *Signal Process.* **84**, 351–365 (2004).
13. T. J. O'shea, T. Roy, and T. C. Clancy, "Over the air deep learning based radio signal classification," *IEEE J. Sel. Top. Signal Process.* **12**, 168–179 (2018).

14. D. Wang, M. Zhang, J. Li, Z. Li, J. Li, C. Song, and X. Chen, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Express* **25**, 17150–17166 (2017).
15. L. Appeltant, M. C. Soriano, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," *Nat. Commun.* **2**, 468 (2011).
16. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," *Nat. Commun.* **4**, 1364 (2013).
17. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," *Opt. Express* **20**, 3241–3249 (2012).
18. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Fast photonic information processing using semiconductor lasers with delayed optical feedback: role of phase dynamics," *Opt. Express* **22**, 8672–8686 (2014).
19. L. Larger, A. Baylón-Fuentes, R. Martinenghi, V. S. Udal'tsov, Y. K. Chembo, and M. Jacquot, "High-speed photonic reservoir computing using a time-delay-based architecture: million words per second classification," *Phys. Rev. X* **7**, 011015 (2017).
20. Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *Sci. Rep.* **2**, 287 (2012).
21. R. Martinenghi, S. Rybalko, M. Jacquot, Y. K. Chembo, and L. Larger, "Photonic nonlinear transient computing with multiple-delay wavelength dynamics," *Phys. Rev. Lett.* **108**, 244101 (2012).
22. F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," *Opt. Express* **20**, 22783–22795 (2012).
23. K. Hicke, M. A. Escalona-Morán, D. Brunner, M. C. Soriano, I. Fischer, and C. R. Mirasso, "Information processing using transient dynamics of semiconductor lasers subject to delayed feedback," *IEEE J. Sel. Top. Quantum Electron.* **19**, 1501610 (2013).
24. A. Dejonckheere, F. Duport, A. Smerieri, L. Fang, J. L. Oudar, M. Haelterman, and S. Massar, "All-optical reservoir computer based on saturation of absorption," *Opt. Express* **22**, 10868–10881 (2014).
25. R. M. Nguimdo, G. Verschaffelt, J. Danckaert, and G. Van der Sande, "Simultaneous computation of two independent tasks using reservoir computing based on a single photonic nonlinear node with optical feedback," *IEEE Trans. Neural Netw. Learn. Syst.* **26**, 3301–3307 (2015).
26. Q. Vinckier, F. Duport, A. Smerieri, K. Vandoorne, P. Bienstman, M. Haelterman, and S. Massar, "High performance photonic reservoir computer based on a coherently driven passive cavity," *Optica* **2**, 438–446 (2015).
27. J. Bueno, D. Brunner, M. C. Soriano, and I. Fischer, "Conditions for reservoir computing performance using semiconductor lasers with delayed optical feedback," *Opt. Express* **25**, 2401–2412 (2017).
28. J. Vatin, D. Rontani, and M. Sciamanna, "Experimental reservoir computing using VCSEL polarization dynamics," *Opt. Express* **27**, 18579–18584 (2019).
29. X. Tan, Y. Hou, Z. Wu, and G. Xia, "Parallel information processing by a reservoir computing system based on a VCSEL subject to double optical feedback and optical injection," *Opt. Express* **27**, 26070–26079 (2019).
30. Y. Hou, G. Xia, W. Yang, D. Wang, E. Jayaprasath, Z. Jiang, C. Hu, and Z. Wu, "Prediction performance of reservoir computing system based on a semiconductor laser subject to double optical feedback and optical injection," *Opt. Express* **26**, 10211–10219 (2018).
31. J. Qin, Q. Zhao, D. Xu, H. Yin, Y. Chang, and D. Huang, "Optical packet header identification utilizing an all-optical feedback chaotic reservoir computing," *Mod. Phys. Lett. B* **30**, 1650199 (2016).
32. J. Qin, Q. Zhao, H. Yin, Y. Jin, and C. Liu, "Numerical simulation and experiment on optical packet header recognition utilizing reservoir computing based on optoelectronic feedback," *IEEE Photon. J.* **9**, 7901311 (2017).
33. A. Argyris, J. Bueno, and I. Fischer, "Photonic machine learning implementation for signal recovery in optical communications," *Sci. Rep.* **8**, 8487 (2018).
34. A. Argyris, J. Bueno, and I. Fischer, "PAM-4 transmission at 1550 nm using photonic reservoir computing post-processing," *IEEE Access* **7**, 37017–37025 (2019).
35. H. Jaeger, "Short term memory in echo state networks," *GMD Rep.* **152**, 60 (2002).
36. A. Wang, Y. Wang, and H. He, "Enhancing the bandwidth of the optical chaotic signal generated by a semiconductor laser with optical feedback," *IEEE Photon. Technol. Lett.* **20**, 1633–1635 (2008).
37. A. Wang, Y. Wang, and J. Wang, "Route to broadband chaos in a chaotic laser diode subject to optical injection," *Opt. Lett.* **34**, 1144–1146 (2009).
38. L. Appeltant, G. Van der Sande, J. Danckaert, and I. Fischer, "Constructing optimized binary masks for reservoir computing with delay systems," *Sci. Rep.* **4**, 3629 (2015).
39. M. C. Soriano, S. Ortín, D. Brunner, L. Larger, C. R. Mirasso, I. Fischer, and L. Pesquera, "Optoelectronic reservoir computing: tackling noise-induced performance degradation," *Opt. Express* **21**, 12–20 (2013).
40. J. Nakayama, K. Kanno, and A. Uchida, "Laser dynamical reservoir computing with consistency: an approach of a chaos mask signal," *Opt. Express* **24**, 8679–8692 (2016).
41. R. Lang and K. Kobayashi, "External optical feedback effects on semiconductor injection laser properties," *IEEE J. Quantum Electron.* **16**, 347–355 (1980).
42. S. Okada, M. Ohzeki, and S. Taguchi, "Efficient partition of integer optimization problems with one-hot encoding," *Sci. Rep.* **9**, 13036 (2019).
43. Q. Xiang, Y. Yang, Q. Zhang, and Y. Yao, "Joint and accurate OSNR estimation and modulation format identification scheme using the feature-based ANN," *IEEE Photon. J.* **11**, 7204211 (2019).
44. Y. Bengio and Y. Grandvalet, "No unbiased estimator of the variance of K-fold cross-validation," *J. Mach. Learn. Res.* **5**, 1089–1105 (2004).
45. P. Li, Q. Cai, J. Zhang, B. Xu, Y. Liu, A. Bogirs, K. A. Shore, and Y. Wang, "Observation of flat chaos generation using an optical feedback multi-mode laser with a band-pass filter," *Opt. Express* **27**, 17859–17867 (2019).
46. A. Zhao, N. Jiang, C. Xue, J. Tang, and K. Qiu, "Wideband complex-enhanced chaos generation using a semiconductor laser subject to delay-interfered self-phase-modulated feedback," *Opt. Express* **27**, 12336–12348 (2019).
47. B. G. Mobasseri, "Digital modulation classification using constellation shape," *Signal Process.* **80**, 251–277 (2000).
48. L. Guesmi, A. M. Ragheb, H. Fathallah, and M. Menif, "Experimental demonstration of simultaneous modulation format/symbol rate identification and optical performance monitoring for coherent optical systems," *J. Lightwave Technol.* **36**, 2230–2239 (2017).
49. W. S. Saif, A. M. Ragheb, H. E. Seleem, T. A. Alshawi, and S. A. Alshebeili, "Modulation format identification in mode division multiplexed optical networks," *IEEE Access* **7**, 156207 (2019).
50. I. Park, I. Fischer, and W. Elsäßer, "Highly nondegenerate four-wave mixing in a tunable dual-mode semiconductor laser," *Appl. Phys. Lett.* **84**, 5189–5191 (2004).

# PHOTONICS Research

## Integrating deep learning to achieve phase compensation for free-space orbital-angular-momentum-encoded quantum key distribution under atmospheric turbulence

XINGYU WANG,<sup>1,2</sup>  TIANYI WU,<sup>2</sup> CHEN DONG,<sup>2,\*</sup> HAONAN ZHU,<sup>1</sup> ZHUODAN ZHU,<sup>3</sup> AND SHANGHONG ZHAO<sup>1</sup>

<sup>1</sup>School of Information and Navigation, Air Force Engineering University, Xi'an 710077, China

<sup>2</sup>Information and Communication College, National University of Defense Technology, Xi'an 710006, China

<sup>3</sup>No. 94782 Unit of PLA, Hangzhou 310021, China

\*Corresponding author: dongchengfk@163.com

Received 9 September 2020; revised 17 November 2020; accepted 20 November 2020; posted 22 November 2020 (Doc. ID 409645); published 13 January 2021

A high-dimensional quantum key distribution (QKD), which adopts degrees of freedom of the orbital angular momentum (OAM) states, is beneficial to realize secure and high-speed QKD. However, the helical phase of a vortex beam that carries OAM is sensitive to the atmospheric turbulence and easily distorted. In this paper, an adaptive compensation method using deep learning technology is developed to improve the performance of OAM-encoded QKD schemes. A convolutional neural network model is first trained to learn the mapping relationship of intensity profiles of inputs and the turbulent phase, and such mapping is used as feedback to control a spatial light modulator to generate a phase screen to correct the distorted vortex beam. Then an OAM-encoded QKD scheme with the capability of real-time phase correction is designed, in which the compensation module only needs to extract the intensity distributions of the Gaussian probe beam and thus ensures that the information encoded on OAM states would not be eavesdropped. The results show that our method can efficiently improve the mode purity of the encoded OAM states and extend the secure distance for the involved QKD protocols in the free-space channel, which is not limited to any specific QKD protocol. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.409645>

### 1. INTRODUCTION

Quantum key distribution (QKD) can provide information theoretic security to share keys between two distant parties [1–3]. Currently, free-space quantum communication has progressed out of laboratories into real-world scenarios [4–8], which paves the way towards global-scale and highly secure quantum communication networks. However, most experiments mainly rely on either the polarization [9,10] or phase [11] of faint laser pulses as information carriers, where the typical two-dimensional encoding scheme limits the capacity of QKD systems due to an intrinsically bounded Hilbert space.

Unlike the limited degree of freedom on polarization or phase states, the orbital angular momentum (OAM) is a high-dimensional encoding scheme for free-space QKD owing to an infinite number of available OAM eigenstates in principle. Recently, Refs. [12–15] implemented OAM-encoded QKD experiments, and Ref. [16] realized OAM-based entanglement distribution over a free-space optical (FSO) channel of more than 143 km. Furthermore, OAM has been experimentally

demonstrated to be rotational invariant in the propagation direction, which can remove the error caused by reference frame misalignment [17]. These outstanding properties make OAM states useful in both classical communication [18–22] and high-dimensional quantum cryptography [23–31].

Nonetheless, it has been shown that the helical phase of a vortex beam that carries OAM is sensitive to the transmission environment and easily distorted [32–34]. Particularly in a free-space optical channel, the atmospheric turbulence resulting from the inhomogeneity of temperature and pressure in the atmosphere will lead to severe wavefront distortion and coherence destruction of the beam, which will directly increase the crosstalk among the adjacent OAM modes [35,36] and further influence the performance of the key rate in OAM-encoded QKD schemes. Some meaningful methods, such as a post-selection of the data or an increase in the mode spacing, which are usually used in OAM-encoded QKD systems to improve the stability of vortex beams against atmospheric turbulence, will lead to a reduction of single-photon gain and coding dimension. In addition, an adaptive optics (AO) system is often

applied in the classical OAM-encoded communications to implement the phase compensation [37–39]. In such scenarios, the compensation accuracy depends on the convergence of algorithms used in AO systems, which requires relatively more processing time. Especially for the satellite-based QKD system with time-varying channel conditions, the AO systems may not be fully adequate for an efficient turbulence correction. Therefore, methods for quickly compensating for the turbulence-distorted vortex beams are still urgently needed for practically applying OAM-encoded QKD systems, especially for satellite-based QKD applications with the limited link durations.

In this paper, we propose a fast and efficient method to adaptively compensate for the atmospheric turbulence effects, which borrows the ideas of deep learning technology. Deep learning technology is one of the most popular artificial intelligence (AI) tools, and it shows extraordinary ability in image recognition [40], optical fields [41], and even parameter optimization for QKD [42,43]. In our proposed method, a convolutional neural network (CNN) model is first designed to predict phase distortion information of atmospheric turbulence. Then CNN combined with common setups is integrated as an accurate and stable phase feedback module to generate a phase screen [44–46] and correct the distorted vortex beam in our free-space OAM-encoded QKD scheme. Our results show that the compensation ability of our method is independent of the turbulence intensity and only relies on the structure of the CNN. In addition, our method can efficiently eliminate the turbulence distortion and improve the mode purity of OAM states compared to the OAM-encoded QKD scheme without compensation, which extends the secure distance for the involved QKD protocols in the free-space channel.

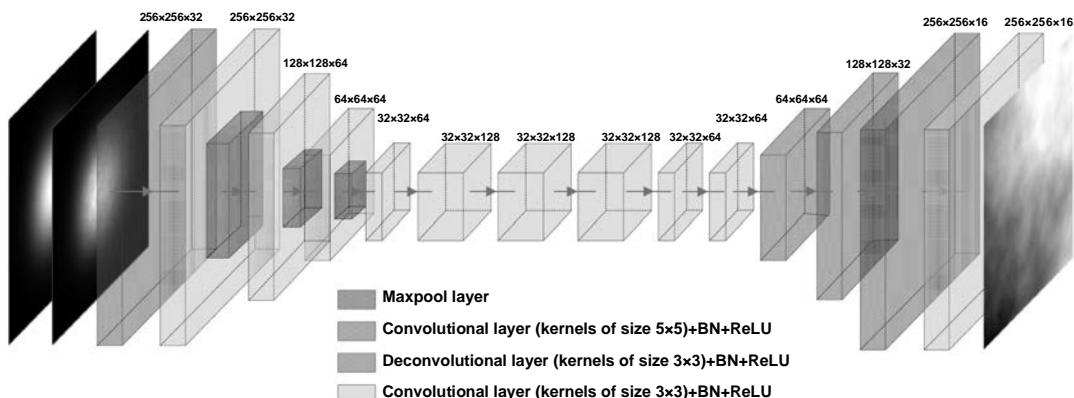
The organization of the article is as follows. In Section 2, a CNN is presented for estimating phase distortion induced by atmospheric turbulence. Then, in Section 3, we describe the components and phase compensation operation of our free-space OAM-encoded QKD scheme and show the improvement of mode purity of OAM states at different situations

to verify the effectiveness of our method. In Section 4, we analyze the performance of the involved OAM-encoded QKD scheme with the decoy-state method. The article is ended in Section 5 with concluding remarks.

## 2. DEEP LEARNING FOR PHASE DISTORTION PREDICTION IN THE TURBULENCE CHANNEL

When the vortex beam transmits through atmospheric turbulence, the refractive index fluctuation causes the helical phase wavefront distortion. In our work, the phase screen produced by the Kolmogorov model [47], which has been widely used to model the phase distortion information over the propagating beam, is adopted to provide quantitative agreement with analytical results. In parallel, in order to simulate the atmospheric turbulence using the random phase screen located in front of the receiver, the propagation of the optical beam through the atmospheric turbulence channel is usually divided into two processes: vacuum transmission and phase modulation by the turbulence phase screen (details in Appendix A). Consequently, we first need to predict the phase screen containing the distortion information induced by the atmospheric turbulence to perform the phase compensation for correcting the distorted vortex beam.

The CNN is composed of multiple-layered structures built from human-brain-like behavior (i.e., neurons), which can learn data with multiple levels of abstraction and can extract high-level abstract features. Based on the universal approximation theorem of a neural network, it is possible to infinitely approximate any complicated but smooth functions on a defined domain with a CNN. Stated differently, it can be applied to accept new incoming data  $x$  and predict the corresponding  $\hat{y}$  after adequate iterations. Here we designed a CNN to learn the phase screen in our atmospheric turbulence prediction work as illustrated in Fig. 1. The CNN structure consists of 15 learned layers, including 12 convolutional layers and 3 deconvolutional layers. The first convolutional layer uses  $5 \times 5$  convolutional kernels, and the other layers use  $3 \times 3$  convolutional kernels.



**Fig. 1.** CNN structure is used to predict the phase distortion in atmospheric turbulence channel. BN, batch normalization; ReLU, rectified linear unit. We consider the encoder-decoder architecture of SegNet [48], since the prediction of the turbulent phase screen belongs to the pixel-by-pixel prediction tasks. The encoder part includes the convolutional layers, pooling layers, etc., where the three maxpool layers make the resolution of the feature map 8 times lower than that of the original image to perform nonlinear upsampling. In the decoder part, three deconvolutional layers are used for upsampling by a factor of 8, so as to keep the size of the original input image consistent. Note that the CNN can be changed according to the needs of specific prediction tasks.

The CNN is trained with a set of training data. This set contains  $N$  mappings  $(x_i^T, y_i^T)$ ,  $i = 1, 2, \dots, N$ . Specifically,  $y_i^T$  is a gray image of the random phase screen produced by the Kolmogorov model in different turbulence intensities, and the input  $x_i^T$  represents the two gray images, i.e., the intensity distribution of the Gaussian probe beam (GPB) affected by turbulence and that without turbulence. Namely, the network can analyze distortion information in the form of images, which can automatically find the characterization of turbulence to generate a predicted phase screen for the subsequent phase compensation.

In our CNN, a loss function mean squared error (MSE) is defined on the output layer between the network's calculated outputs  $\hat{y}_i = f(x_i^T, w)$  on a set of input data  $x_i^T$ , versus the set of desired output  $y_i^T$ . The loss function MSE takes the form

$$\text{MSE} = \frac{1}{N} \sum_i^N [f(x_i^T, w) - Y]^2. \quad (1)$$

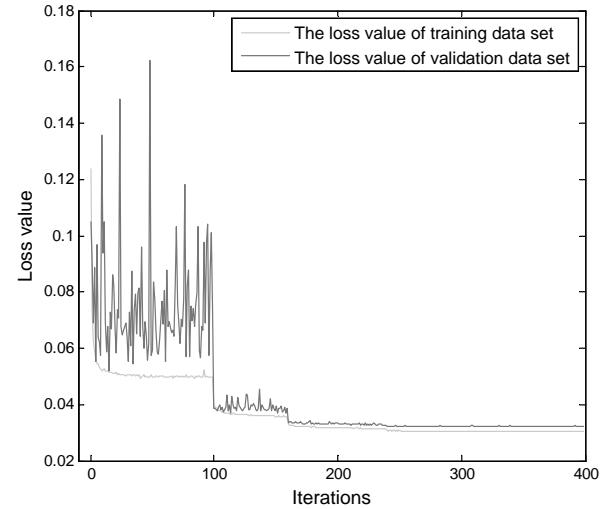
Here we take a linear combination of inputs  $x_i^T$ , with weight  $w$  and constant factor  $b$ , and the activation in each neuron can be expressed as

$$\begin{aligned} l_1[f(x_i^T, w), y_i] &= (y_i - \hat{y}_i)^2, \\ l_2[f(x_i^T, w), y_i] &= \text{ReLU}(y_i - \hat{y}_i - b), \\ l_3[f(x_i^T, w), y_i] &= \text{ReLU}(\hat{y}_i - y_i - b), \end{aligned} \quad (2)$$

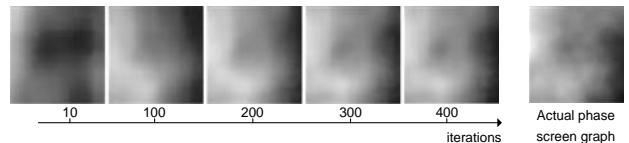
where ReLU represents rectified linear unit function [i.e.,  $\max(0, x)$ ]. We use the backpropagation algorithm [49] to quickly compute the partial derivatives of the loss function to the internal weight  $w$  in our CNN. Simultaneously, we adjust the weight  $w$  accordingly via the mini-batch gradient descent (MBGD) algorithm [50] to minimize the loss function and make  $\hat{y}_i$  approach  $y_i^T$  as much as possible. Finally, by substituting Eq. (2) into Eq. (1), the loss function MSE between the estimated value and actual value of phase information is calculated as follows:

$$\begin{aligned} \text{MSE} &= \frac{1}{N} \sum_{i=1}^N \{l_1[f(x_i^T, w), y_i] + l_2[f(x_i^T, w), y_i] \\ &\quad + l_3[f(x_i^T, w), y_i]\} \\ &= \frac{1}{N} \sum_{i=1}^N [(y_i - \hat{y}_i)^2 + \text{ReLU}(y_i - \hat{y}_i - b) \\ &\quad + \text{ReLU}(\hat{y}_i - y_i - b)]. \end{aligned} \quad (3)$$

To achieve the trained CNN with a better generalization ability to prevent overfitting, we use an amount of training data to decrease the training error and the model complexity. Here we use 84,000 GPB intensity images with specified turbulence (i.e., the atmospheric refractive index structure constant  $C_n^2$ ) ranging from  $[10^{-15}, 10^{-13}]$ , where the size of each image is fixed at  $256 \times 256$  and the gray values are normalized into  $[0, 1]$ . In detail, 80,000 of them are used as the training data, 2000 are for the test data, and the remaining 2000 are for the validation data. Figure 2 shows the loss function, where the yellow and blue curves represent the loss value of the validation data set and training data set, respectively. The loss value of the



**Fig. 2.** Loss values of the validation data set and training data set against the iterations of the algorithm. Here we dynamically tune the learning rate at every 10 epochs. As the number of iterations increases, the learning rate is gradually reduced in the process of approaching the optimal solution, which shows drops of the loss value at 100 iterations.



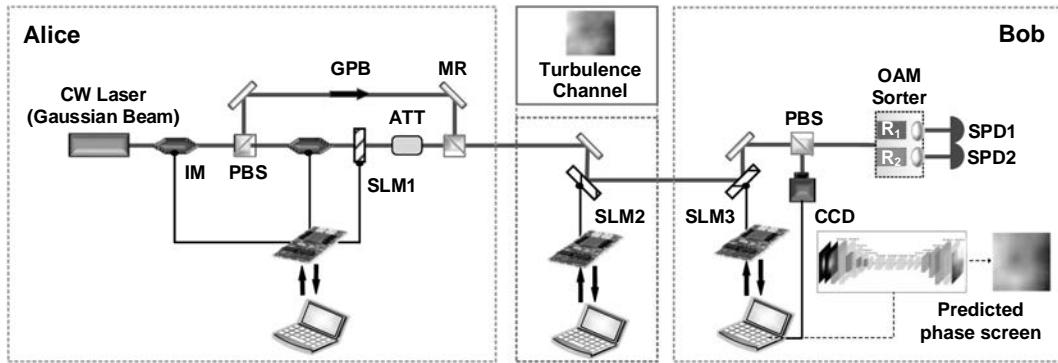
**Fig. 3.** Atmospheric turbulence phase graphs predicted by the CNN different training iterations. The five images from left to right represent the phase screens that are predicted by the validation data set at 10, 100, 200, 300, and 400 iterations, and the rightmost one is the one corresponding to the actual phase screen.

validation data set is very close to that of training data set, and both converge to be stable. After 400 iterations, the loss value of the validation data set dropped to 0.03, which demonstrated that the CNN at this time has a good performance of prediction on turbulence and no overfitting occurred.

To reflect the trend of the loss curve more intuitively, Fig. 3 presents the phase images predicted by our CNN. From the images, we can see that when the number of iterations is sufficient, the predicted phase images are almost the same as the real phase screens. In detail, the time required for the iterations and processing is about 0.1 ms per image by using an i9-8950 CPU, which is much lower than the turbulent freezing time (10 ms) [51], and thus it brings advantages to optimize the performance of the OAM-encoded QKD scheme.

### 3. DEEP-LEARNING-BASED PHASE COMPENSATION IN THE FREE-SPACE OAM-ENCODED QKD SCHEME

We use the trained CNN for phase compensation in our OAM-encoded QKD scheme. Figure 4 shows the schematic diagram of an OAM-encoded QKD system with decoy-state method, which contains the real-time atmospheric turbulence



**Fig. 4.** Schematic diagram of an OAM-encoded QKD with decoy-state method. CW laser, continuous-wave laser; GPB, Gaussian probe beam; IM, intensity modulator; PBS, polarization beam splitter; SLM, spatial light modulator; ATT, attenuator; MR: mirror; CCD, charge-coupled device; SPD, single-photon detector. At Alice's side, the signal pulses (a CW laser beam modulated by an IM) are separated to two parts by a PBS. The  $x$ -polarized GPB is modulated to produce a decoy state and encoded to load the spatially helical phase wavefront (or OAM) by an IM and SLM1, respectively. The  $y$ -polarized GPB is then combined with the encoded OAM signal pulses by another PBS and transmitted together through turbulence channel simulated by SLM2. Both the  $y$ -polarized GPB and the encoded OAM signal pulses will get phase distortion when transmitted through the simulated turbulence channel. The phase distortion of the  $y$ -polarized GPB is detected by a CCD, and the corresponding reverse phase screen will be produced by SLM3 to compensate for the phase distortion of the OAM-encoded signal pulses. After compensation, the OAM-encoded signal pulses will be classified by the OAM sorter and received by different SPDs.

compensation feedback at Bob's side. In detail, at Alice's side, a Gaussian beam from a continuous-wave (CW) laser (operating at 1550 nm) is modulated into a pulse train with a 1 MHz repetition rate. A polarization beam splitter (PBS) is utilized to separate the pulses into two subpulses, where the signal pulses are linearly polarized in the  $x$  direction and the Gaussian probe pulses are linearly polarized in the  $y$  direction. In practice, it is usually difficult to create a perfect vacuum state in the decoy-state QKD experiments [52]. To realize the preparation of decoy states, an intensity modulator is used to modulate the signal pulses into different intensities. When the gain of a state is modulated to be close (a finite extinction ratio below 30 dB) to the dark count rate of our single-photon detector, the state will be regarded as the “vacuum state.” After that, the signal pulses are converted by spatial light modulator (SLMs) to carry OAM-encoded qubit signal and then attenuated to a single-photon level signal by an attenuator. By using a polarizing beam splitter (PBS), the quantum signals together with the Gaussian probe pulses are sent to Bob through a simulated free-space channel.

In the free-space channel, the equivalent phase screens are produced by SLM2 under the Kolmogorov model. At Bob's side, the compensation feedback is used to perform phase compensation, among which SLM3 is only sensitive to  $x$  polarization and a CCD is utilized to extract the intensity distribution of Gaussian probe pulses, which is used as the input of a laptop PC with our CNN to predict a real-time atmospheric turbulence phase screen. Thus, by getting a reversed phase of the predicted phase screen on SLM3, the phase distortion of the only vortex pulses induced by atmospheric turbulence can be restored. After these operations, the OAM states are differentiated by employing the OAM sorter [12], which consists of two static optical transformations [unwrapper (R1) and phase corrector (R2) shown in Fig. 4] and a lens. Finally, two single-photon detectors are adopted to detect the results.

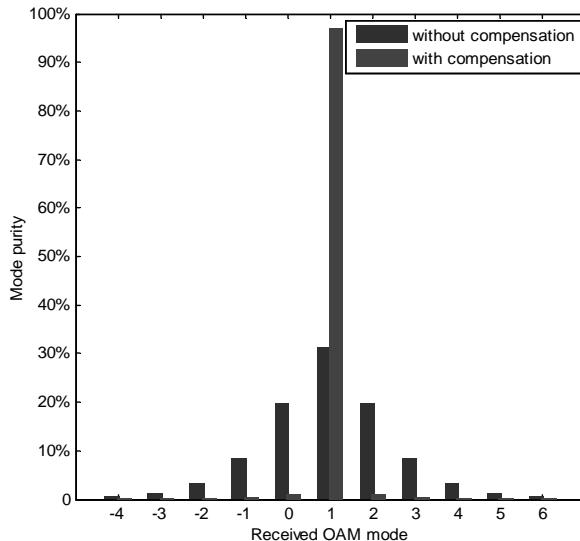
Note that the measurement probabilities of the desired OAM state relate directly to the probabilities of obtaining correct or incorrect measurements of a transmitted “symbol.” Therefore, for quantitative analysis of the correction effect in terms of the measurement probability, we now need to compare the mode purity of the OAM states before and after compensation. The mode purity, which can be regarded as the normalized power, is an important parameter describing the crosstalk between different spatial modes. Specifically, we assume that the receiver may measure a photon with the adjacent OAM mode  $l' = l_0 + \Delta l$ ,  $\Delta l = 0, 1, 2, \dots$ . Thus, the mode purity of OAM states can be estimated by taking ratio  $s_\Delta = P_\Delta / P$  between the power contained in each OAM mode and the total power collected by the receiver. Without loss of generality, we are only interested in the ensemble average of this quantity (i.e.,  $\langle s_\Delta \rangle$ ). Meanwhile, the OAM topological charge is set as  $l_0 = 1$ . For Kolmogorov turbulence theory, the mode purity of an OAM state propagating through turbulence can be written as [34]

$$\langle s_\Delta \rangle = \frac{1}{\pi} \int_0^1 \frac{r}{R} d\frac{r}{R} \int_0^{2\pi} \exp \left[ -6.88 \times 2^{2/3} \left( \frac{r}{r_0} \right)^{5/3} \left| \sin \frac{\Delta\theta}{2} \right|^{5/3} \right] \times \exp(-i\Delta\theta/\Delta\theta) d\Delta\theta, \quad (4)$$

where  $r$  and  $\theta$  are the radial coordinates and the azimuthal coordinates, respectively;  $r_0$  is the Fried parameter [10]; and  $R$  is the radius of a receiving aperture. By substituting  $\Delta l = 0$  into Eq. (4), we can get the mode purity of the receiving initial OAM state  $l_0 = 1$  without phase compensation:

$$\langle s_0 \rangle = \frac{1}{\pi} \int_0^1 \frac{r}{R} d\frac{r}{R} \int_0^{2\pi} \exp \left[ -6.88 \times 2^{2/3} \left( \frac{r}{r_0} \right)^{5/3} \left| \sin \frac{\Delta\theta}{2} \right|^{5/3} \right] d\Delta\theta. \quad (5)$$

On the other hand, as discussed in Ref. [53], the effect of our phase compensation feedback on OAM modes can also be



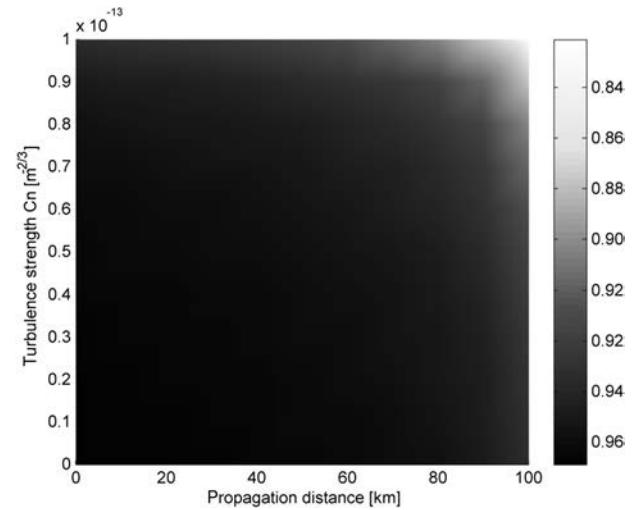
**Fig. 5.** Mode purity of the received OAM states with and without compensation.

described by the Strehl ratio [54], whose calculation is based on the root mean square (RMS) of the residual wavefront aberration. In our case, the correction errors caused by the system are all attributed to the accuracy of the algorithm in our CNN. Therefore, combined with the performance indices of our CNN, i.e., the MSE value  $\chi = 1$ , the mode purity of the receiving initial OAM state with phase compensation can be estimated:

$$\langle s_0 \rangle' = \exp(-\chi). \quad (6)$$

By numerical simulation, we present the mode purity of the received OAM states with and without compensation in Fig. 5. Here the beam wavelength  $\lambda = 1550$  nm, the waist radius  $\omega_0 = 0.002$  m, and the radius of receiving aperture  $R = 0.075$  m. The propagation distance  $\Delta z$  is set as 50 m, and the atmospheric refractive index structure constant  $C_n^2 = 10^{-14} \text{ m}^{-2/3}$ . It is clear from Fig. 5 that the probabilities for an OAM mode scattering to adjacent OAM modes are greatly reduced, and the mode purity of the initial OAM state  $l_0 = 1$  is significantly improved from 32.5% to 97.1%.

To show the applicability of our method in various environments, we update the training data set for our CNN, where the ranges of  $\Delta z$  and  $C_n^2$  are set as  $[1 \times 10^3, 1 \times 10^5]$  m and  $[10^{-14}, 10^{-13}] \text{ m}^{-2/3}$ , respectively. In particular, we generate 70,000 GPB images every 10 km to form the final data set, i.e., a total of  $10 \times 70,000$  pieces of data set to train the CNN, where the entire process from generating the data set to training the CNN took about 6 h. After retraining, we show the mode purity of OAM state  $l_0 = 1$  at different situations. As shown in Fig. 6, the mode purity can always be well recovered under a wide range of parameter combinations of the intensity of turbulence and the distance. Although the prediction effect under strong turbulence or long distance is reduced (some pixels of the phase screen are out of the gray value range, which leads to the reduction of effective information during the normalization process), the mode purity values in different situations are still



**Fig. 6.** Mode purity of the received OAM state in the different situations.

improved to above 0.82. This is because the most key working mechanism of the proposed CNN is to find the mapping relationship between inputs and outputs. Next, we will discuss the key rate performance optimization of the OAM-encoded QKD system.

#### 4. PERFORMANCE OPTIMIZATION OF THE FREE-SPACE OAM-ENCODED QKD

To further evaluate the performance of the deep-learning-based phase compensation in improving the performance of the OAM-encoded QKD scheme with the decoy-state method, we discuss the secret key rate of the involved system. For simplicity and generality, a two-dimensional OAM-QKD protocol is discussed, where the OAM basis and superposition basis (SUP) [12] are both used in the OAM-QKD protocol. In such a decoy-state QKD, Alice randomly chooses an OAM basis ( $| -l_0 \rangle | l_0 \rangle$ ) or the SUP basis consisting of  $\frac{1}{\sqrt{2}}(| l_0 \rangle + e^{inx}| -l_0 \rangle)$ , where  $n = 0, 1$  is the encoding bit. Subsequently, Bob uses the results with OAM basis as the key bits and the results with SUP basis as the testing bits, and then they testify the security of the key distribution. After error correction and privacy amplification, they can distill a secure key.

We follow the decoy-state QKD theory from Ref. [55]. Using the Gottesman–Lo–Lutkenhaus–Preskill (GLLP) formula [56], the secure key rate of decoy-state OAM-QKD in the asymptotic case is given by

$$R_{\text{GLLP}} = q \{-f_e(E_\mu^{\text{OAM}})Q_\mu^{\text{OAM}}h_2(E_\mu^{\text{OAM}}) + \mu e^{-\mu} Y_1^{\text{OAM}}[1 - h_2(e_1^{\text{SUP}})]\}, \quad (7)$$

where  $q$  depends on the implementation (1/2 for the BB84 protocol due to the fact that half of the time Alice and Bob disagree with the bases, and if one uses the efficient BB84 protocol,  $q \approx 1$ );  $f_e$  is the error correction inefficiency function,  $\mu$  is the intensity of the signal state, and  $h_2$  is the binary entropy function.  $Q_\mu^{\text{OAM}}$  and  $E_\mu^{\text{OAM}}$  are the total gain and error rate

**Table 1.** List of Parameters Used in the Simulations

Experimental Parameters	Value
Radius of the receiving aperture $R_r$ (cm)	7.5
Signal wavelength $\lambda$ (nm)	1550
Refraction structure constant $C_n^2$	$10^{-14}$
Error rate of dark count $e_0$	0.5
Misalignment error $e_d$	0.015
Error correction inefficiency $f_e$	1.22
Background dark count rate $Y_0$	$3 \times 10^{-6}$
Detection efficiency of the SPD $\eta_d$	50%
Finite size of data $N$	$10^{14}$

under the OAM basis, which can be theoretically calculated by the standard channel model [57] using a transmittance of  $\eta$ .

Here the transmittance consists of the probabilities of obtaining correct and incorrect initial OAM states. As discussed in the below section, the probability of obtaining a correct state can be expressed as

$$\eta_0 = e^{-\beta L} \langle s_0 \rangle', \quad (8)$$

where  $\beta$  is the link attenuation coefficient. Correspondingly, the other one, the transmission error probability, is  $\bar{\eta}_0 = 1 - \eta_0$ . Thus,  $Q_\mu^{\text{OAM}}$  can be calculated by

$$Q_\mu^{\text{OAM}} = Y_0 + 1 - e^{-\mu \cdot \eta_d}. \quad (9)$$

Similarly, the gain  $Q_\mu^{\text{SUP}}$  in the SUP basis can also be estimated in a similar way. Therefore,  $Q_\mu^{\text{SUP}} = Q_\mu^{\text{OAM}}$ . Note that, considering the errors are induced by misalignment and crosstalk between states,  $E_\mu^{\text{OAM}}$  in the OAM basis can be given as

$$E_\mu^{\text{OAM}} = \frac{e_0 Y_0 + t \cdot e_d (Q_\mu^{\text{OAM}} - Y_0)}{Q_\mu^{\text{OAM}}}, \quad (10)$$

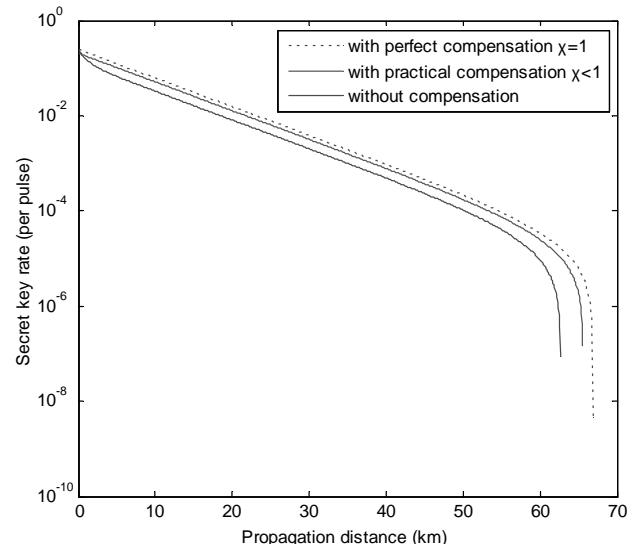
where  $e_0$  is the error rate of the dark count,  $Y_0$  is the background count rate for the detector, and  $t = \bar{\eta}_0$  is the crosstalk probability. Similarly, the total errors  $E_\mu^{\text{SUP}} = E_\mu^{\text{OAM}}$ . By using decoy-state technique to combine  $Q_\mu^{\text{OAM}}$  and  $E_\mu^{\text{SUP}}$  for different intensities, we can estimate the yield and the error rate of the single-photon states  $Y_1^{\text{OAM}}$  and  $e_1^{\text{SUP}}$ , which can be written as

$$Y_1^{\text{OAM}} = \frac{\mu}{\mu v - v^2} \left( Q_\nu^{\text{OAM}} e^\nu - Q_\mu^{\text{OAM}} e^\mu \frac{v^2}{\mu^2} - e^\mu \frac{\mu^2 - v^2}{e_0 \mu^2} \right), \quad (11)$$

$$e_1^{\text{SUP}} = \frac{E_\nu^{\text{SUP}} Q_\nu^{\text{SUP}} e^\nu - e_0 Y_0}{Y_1^{\text{SUP}} v}. \quad (12)$$

For a fair comparison, we consider the same free-space channel model proposed in Ref. [12], where the link attenuation  $\beta$  is set to  $1.38 \times 10^{-4}$  (0.6 dB/km). The radius of receiving aperture  $R_r$  is 7.5 cm, and the refraction structure constant  $C_n^2$  is fixed at  $10^{-14} \text{ m}^{-2/3}$ . Other experimental parameters are taken directly from Ref. [7] as listed in Table 1.

In Fig. 7, we compare the secret key rates versus propagation distance for the situation with perfect phase compensation (i.e.,  $\chi = 1$ ), the practical situation with phase compensation feedback based on the learned CNN (i.e., the value of  $\chi$  is



**Fig. 7.** Secret key rate performance of the OAM-encoded QKD versus propagation distance for different schemes. From top to bottom, the curves represent the secret key rates in the perfect phase compensation (red dotted line,  $\chi = 1$ ), the practical situation with the phase compensation feedback supported by the learned CNN (blue solid line,  $\chi < 1$ ), and the situation without phase compensation (green dashed line,  $\chi = 0$ ). Note that we fix the signal and decoy intensities to  $\mu = 0.3$ ,  $\nu = 0.05$ . Therefore, by scanning through the decoy-state intensities and probabilities, the key rate can be further optimized.

dependent on the transmission distance), and the situation without phase compensation. As shown in Fig. 7, our method offers a higher key rate than that of no compensation. Even if the learned CNN is not perfect, the curves of the practical situation for the secret key rates are quite close to the perfect phase compensation, where the situation of perfect phase compensation can be regarded as having no turbulence effects. This means that our proposed method can be applied in the free-space OAM-encoded QKD system to mitigate the turbulence phase distortion.

## 5. CONCLUSIONS

In this work, considering that the helical phase wavefront induced by the turbulence distortion has significant influence on the performance of a free-space OAM-encoded QKD scheme with the decoy-state method, we integrate the deep learning technology to achieve phase compensation to optimize the involved system. Here a well-trained CNN is designed to predict an equivalent turbulent phase screen. Combined with the commonly used optical devices, fast and accurate phase compensation feedback can be achieved in our experimental scheme. Furthermore, we use the mean-squared error (MSE) between the predicted phase values and original phase values as a performance index for assessing our system. The results show that our scheme can effectively increase the key rate owing to improving the mode purity of OAM states. Our proposed work can provide an efficient solution for a free-space high-dimensional QKD using OAM states.

Note that, considering the time-varied atmospheric turbulence, we use a charge-coupled device (CCD) to extract the real-time turbulence distortions from the intensity distribution of the Gaussian probe beam, which has no coding information. By employing an SLM, a generated compensated phase screen can be directly loaded on the OAM-encoded beams, and thus there is no need to monitor the OAM-encoded signals. In experiments, limited by the response speed of devices such as the SLM, our experimental scheme may degrade the compensation effect, but it can reduce the number of qubits used to estimate channel conditions. Moreover, attacks due to the loopholes of the phase compensation module can also be resisted. In addition, placing the phase compensation module on the measurement side [i.e., measurement-device-independent (MDI)-QKD protocol] can further ensure its safety. Therefore, the experimental and security analysis of the MDI-QKD system integrated with our method will be the subject of our future studies.

## APPENDIX A

When the vortex beam transmits through atmospheric turbulence, the refractive index fluctuation causes helical phase wave-front distortion. To simulate the atmospheric turbulence, we model the atmospheric channel using a phase screen located in front of the receiver. Thus, the propagation of the optical beam through the atmospheric turbulence channel with a length  $z + \Delta z$  can be divided into two processes: a vacuum transmission channel with a length  $z$  and a phase modulation process by the turbulence phase screen with a width  $\Delta z$ . The phase screen method involves the filtering of a complex Gaussian random field using the phase's power spectral density (PSD) function of the atmospheric turbulence, where the phase's PSD function  $\Phi(k_x, k_y)$  can be calculated by the Kolmogorov model:

$$\Phi(k_x, k_y) = 2\pi k_0^2 \Delta z \cdot 0.033 C_n^2 \sqrt{k_x^2 + k_y^2}^{-11/3}, \quad (\text{A1})$$

where  $k_0$  is the wavenumber and  $k_x$  and  $k_y$  are the components of  $k_0$  in the  $x$ -axis and  $y$ -axis directions, respectively.  $C_n^2$  is the refractive index structure constant, representing the turbulence intensity. Subsequently, the variance of the phase spectrum can be written as

$$\sigma^2(k_x, k_y) = \left( \frac{2\pi}{N\Delta L} \right)^2 \Phi(k_x, k_y), \quad (\text{A2})$$

where  $\Delta L$  is the grid spacing. Therefore, using the fast Fourier transform (FFT), realization of the corresponding phase screen in the time domain can be represented as

$$\phi(x, y) = \text{FFT}[\mathbf{C} \cdot \sigma(k_x, k_y)]. \quad (\text{A3})$$

Here  $\phi(x, y)$  is an  $N \times N$  phase screen, and  $\mathbf{C}$  is an  $N \times N$  dimensional complex Gaussian random number array with a variance of 1. According to the diffraction angular spectrum theory, when the vortex beam  $E(z, x, y)$  propagates in atmospheric turbulence over  $\Delta z$ , the optical field at  $z + \Delta z$  can be approximately described by

$$E(z + \Delta z, x, y) = \text{FFT}^{-1}\{\exp(iA\Delta z) \cdot \text{FFT}\{\exp[i\phi(x, y)]\} \cdot E(z, x, y)\}, \quad (\text{A4})$$

where  $\exp(iA\Delta z)$  is the transfer function of Fresnel propagation.

In our work, we choose one of the vortex beam forms, i.e., a Laguerre–Gaussian (LG) beam, making OAM eigenstates correspond to the LG mode set. Here we use the intensity and phase profiles of an LG01 beam after 500 m atmospheric propagation with different  $C_n^2$ . Here the beam waist radius  $w_0$  is set as 0.002 m. The phase screen is set with a numerical grid of  $N \times N = 512 \times 512$  points and a grid spacing of  $\Delta L = 5$  mm.

**Funding.** National Natural Science Foundation of China (11704412); Key Research and Development Program of Shaanxi (2019ZDLGY09-01); Innovative Talents Promotion Plan in Shaanxi Province (2020KJXX-011); National University of Defense Technology (19-QNCXJ-009).

**Disclosures.** The authors declare that there are no conflicts of interest related to this paper.

## REFERENCES

1. N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden, "Quantum cryptography," *Rev. Mod. Phys.* **74**, 145–195 (2002).
2. C. H. Bennett, G. Brassard, and N. D. Mermin, "Quantum cryptography without Bell's theorem," *Phys. Rev. Lett.* **68**, 557–559 (1992).
3. D. Mayers, "Unconditional security in quantum cryptography," *J. ACM* **48**, 351–406 (2001).
4. W. T. Buttler, R. J. Hughes, P. G. Kwiat, S. K. Lamoreaux, G. G. Luther, G. L. Morgan, J. E. Nordholt, C. G. Peterson, and C. M. Simmons, "Practical free-space quantum key distribution over 1 km," *Phys. Rev. Lett.* **81**, 3283–3286 (1998).
5. C. Z. Peng, T. Yang, X. H. Bao, J. Zhang, X. M. Jin, F. Y. Feng, B. Yang, J. Yang, J. Yin, Q. Zhang, N. Li, B. L. Tian, and J. W. Pan, "Experimental free-space distribution of entangled photon pairs over 13 km: towards satellite-based global quantum communication," *Phys. Rev. Lett.* **94**, 150501 (2005).
6. J. Yin, J. G. Ren, H. Lu, Y. Cao, H. L. Yong, Y. P. Wu, C. Liu, S. K. Liao, F. Zhou, Y. Jiang, X. D. Cai, P. Xu, G. S. Pan, J. J. Jia, Y. M. Huang, H. Yin, J. Y. Wang, Y. A. Chen, C. Z. Peng, and J. W. Pan, "Quantum teleportation and entanglement distribution over 100-kilometre free-space channels," *Nature* **488**, 185–188 (2012).
7. T. Schmitt-Manderbach, H. Weier, M. Fürst, R. Ursin, F. Tiefenbacher, T. Scheidl, J. Perdigues, Z. Sodnik, C. Kurtsiefer, J. G. Rarity, A. Zeilinger, and H. Weinfurter, "Experimental demonstration of free-space decoy-state quantum key distribution over 144 km," *Phys. Rev. Lett.* **98**, 010504 (2007).
8. S. K. Liao, W.-Q. Cai, W.-Y. Liu, L. Zhang, Y. Li, J.-G. Ren, J. Yin, Q. Shen, Y. Cao, Z.-P. Li, F.-Z. Li, X.-W. Chen, L.-H. Sun, J.-J. Jia, J.-C. Wu, X.-J. Jiang, J.-F. Wang, Y.-M. Huang, Q. Wang, Y.-L. Zhou, L. Deng, T. Xi, L. Ma, T. Hu, Q. Zhang, Y.-A. Chen, N.-L. Liu, X.-B. Wang, Z.-C. Zhu, C.-Y. Lu, R. Shu, C.-Z. Peng, J.-Y. Wang, and J.-W. Pan, "Satellite-to-ground quantum key distribution," *Nature* **549**, 43–47 (2017).
9. A. Carrasco-Casado, H. Kunimori, H. Takenaka, T. Kubo-Oka, M. Akioka, T. Fuse, Y. Koyama, D. Kolev, Y. Munemasa, and M. Toyoshima, "LEO-to-ground polarization measurements aiming for space QKD using small optical transponder (SOTA)," *Opt. Express* **24**, 12254–12266 (2016).
10. G. Vallone, D. Daniele, M. Tomasin, M. Schiavon, F. Vedovato, D. Bacco, S. Gaiarin, G. Bianco, V. Luceri, and P. Villoresi, "Satellite quantum communication towards GEO distances," *Proc. SPIE* **9900**, 99000J (2016).

11. C. Zhou, W. S. Bao, W. Chen, and H. W. Li, "Phase-encoded measurement-device-independent quantum key distribution with practical spontaneous-parametric-down-conversion sources," *Phys. Rev. A* **88**, 052333 (2013).
12. L. Wang, S. M. Zhao, L. Y. Gong, and W. W. Cheng, "Free-space measurement-device-independent quantum-key-distribution protocol using decoy states with orbital angular momentum," *Chin. Phys. B* **24**, 120307 (2015).
13. M. Mafu, A. Dudley, S. Goyal, D. Giovannini, M. McLaren, M. J. Padgett, T. Konrad, F. Petruccione, N. Lütkenhaus, and A. Forbes, "Higher-dimensional orbital-angular-momentum-based quantum key distribution with mutually unbiased bases," *Phys. Rev. A* **88**, 032305 (2013).
14. S. Zhao, W. Li, Y. Shen, Y. Yu, X. Han, H. Zeng, M. Cai, T. Qian, S. Wang, Z. Wang, Y. Xiao, and Y. Gu, "Experimental investigation of quantum key distribution over a water channel," *Appl. Opt.* **58**, 3902–3907 (2019).
15. Q. H. Lai, M. X. Luo, C. Zhan, J. Pieprzyk, and M. A. Orgun, "An improved coding method of quantum key distribution protocol based on Fibonacci-valued OAM entangled states," *Phys. Lett. A* **381**, 2922–2926 (2017).
16. M. Krenn, J. Handsteiner, M. Fink, R. Fickler, R. Ursin, M. Malik, and A. Zeilinger, "Twisted light transmission over 143 km," *Proc. Natl. Acad. Sci. USA* **113**, 13648–13653 (2016).
17. J. L. Liu and C. Wang, "Six-State quantum key distribution using photons with orbital angular momentum," *Phys. Rev. A* **27**, 110303 (2010).
18. X. Zhang, Y. He, Y. Cai, M. Su, X. Zhou, Y. Chen, S. Chen, Y. Xiang, L. Chen, C. Su, Y. Li, and D. Fan, "Coherent separation detection for orbital angular momentum multiplexing in free-space optical communications," *IEEE Photon. J.* **9**, 7903811 (2017).
19. J. Wang, J. Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, and A. E. Willner, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nat. Photonics* **6**, 488–496 (2012).
20. A. E. Willner, H. Huang, Y. Yan, Y. Ren, N. Ahmed, G. Xie, C. Bao, L. Li, Y. Cao, Z. Zhao, J. Wang, M. P. J. Lavery, M. Tur, S. Ramachandran, A. F. Molisch, N. Ashrafi, and S. Ashrafi, "Optical communications using orbital angular momentum beams," *Adv. Opt. Photon.* **7**, 66–106 (2015).
21. Y. Zhang, P. Wang, L. Guo, W. Wang, and H. Tian, "Performance analysis of an OAM multiplexing-based MIMO FSO system over atmospheric turbulence using space-time coding with channel estimation," *Opt. Express* **25**, 19995–20011 (2017).
22. M. Krenn, R. Fickler, M. Fink, J. Handsteiner, M. Malik, T. Scheidl, R. Ursin, and A. Zeilinger, "Communication with spatially modulated light through turbulent air across Vienna," *New J. Phys.* **16**, 113028 (2014).
23. M. Malik, M. O'sullivan, B. Rodenburg, M. Mirhosseini, J. Leach, M. Lavery, M. J. Padgett, and R. W. Boyd, "Influence of atmospheric turbulence on optical communications using orbital angular momentum for encoding," *Opt. Express* **20**, 13195–13200 (2012).
24. F. Bouchard, K. Heshami, D. England, R. Fickler, R. W. Boyd, B.-G. Englert, L. L. Sánchez-Soto, and E. Karimi, "Experimental investigation of high-dimensional quantum key distribution protocols with twisted photons," *Quantum* **2**, 111 (2018).
25. G. Vallone, V. D'Ambrosio, A. Sponselli, S. Slussarenko, L. Marrucci, F. Sciarrino, and P. Villoresi, "Free-space quantum key distribution by rotation-invariant twisted photons," *Phys. Rev. Lett.* **113**, 060503 (2014).
26. J. P. Zhao, Y. Y. Zhou, B. Braverman, C. Liu, K. Pang, N. Steinhoff, G. Tyler, A. Willner, and R. Boyd, "Performance of real-time adaptive optics compensation in a turbulent channel with high-dimensional spatial-mode encoding," *Opt. Express* **28**, 15376–15391 (2020).
27. A. Sit, F. Bouchard, R. Fickler, J. Gagnon-Bischoff, H. Larocque, K. Heshami, D. Elser, C. Peuntinger, K. Günthner, B. Heim, C. Marquardt, G. Leuchs, R. Boyd, and E. Karimi, "High-dimensional intracity quantum cryptography with structured photons," *Optica* **4**, 1006–1010 (2017).
28. C. Daniele, D. Bacco, B. D. Lio, K. Ingerslev, and L. K. Oxenløwe, "Orbital angular momentum states enabling fiber-based high-dimensional quantum communication," *Phys. Rev. Appl.* **11**, 064058 (2019).
29. C. Daniele, B. D. Lio, D. Bacco, and L. K. Oxenløwe, "High-dimensional quantum communication: benefits, progress, and future challenges," *Adv. Quantum Technol.* **2**, 1900038 (2019).
30. F. Bouchard, A. Sit, F. Hufnagel, A. Abbas, Y. Zhang, K. Heshami, R. Fickler, C. Marquardt, G. Leuchs, R. W. Boyd, and E. Karimi, "Quantum cryptography with twisted photons through an outdoor underwater channel," *Opt. Express* **26**, 22563–22573 (2018).
31. F. Hufnagel, A. Sit, F. Grenapin, F. Bouchard, K. Heshami, D. England, Y. Zhang, B. J. Sussman, R. W. Boyd, G. Leuchs, and E. Karimi, "Characterization of an underwater channel for quantum communications in the Ottawa River," *Opt. Express* **27**, 26346–26354 (2019).
32. N. D. Leonhard, V.-N. Shatokhin, and A. Buchleitner, "Universal entanglement decay of photonic orbital-angular-momentum qubit states in atmospheric turbulence," *Phys. Rev. A* **91**, 012345 (2015).
33. Y. X. Ren, H. Huang, G. D. Xie, N. Ahmed, Y. Yan, B. I. Erkmen, N. Chandrasekaran, M. P. J. Lavery, N. K. Steinhoff, M. Tur, S. Dolinar, M. Neifeld, M. J. Padgett, R. W. Boyd, J. H. Shapiro, and A. E. Willner, "Atmospheric turbulence effects on the performance of a free space optical link employing orbital angular momentum multiplexing," *Opt. Lett.* **38**, 4062–4065 (2013).
34. X. Y. Wang, S. H. Zhao, C. Dong, Z. D. Zhu, and W. Y. Gu, "Orbital angular momentum-encoded measurement device independent quantum key distribution under atmospheric turbulence," *Quantum Inf. Process.* **18**, 304 (2019).
35. C. Paterson, "Atmospheric turbulence and orbital angular momentum of single photons for optical communication," *Phys. Rev. Lett.* **94**, 153901 (2005).
36. G. A. Tyler and R. W. Boyd, "Influence of atmospheric turbulence on the propagation of quantum states of light carrying orbital angular momentum," *Opt. Lett.* **34**, 142–144 (2009).
37. M. Li, M. Cvjetić, Y. Takashima, and Z. Yu, "Evaluation of channel capacities of OAM-based FSO link with real-time wavefront correction by adaptive optics," *Opt. Express* **22**, 31337–31346 (2014).
38. X. Yin, H. Chang, X. Cui, J. Ma, Y. Wang, G. Wu, L. Zhang, and X. Xin, "Adaptive turbulence compensation with a hybrid input-output algorithm in orbital angular momentum-based free-space optical communication," *Appl. Opt.* **57**, 7644–7650 (2018).
39. H. Chang, X. Yin, X. Cui, X. Chen, Y. Su, J. Ma, Y. Wang, L. Zhang, and X. Xin, "Performance analysis of adaptive optics with a phase retrieval algorithm in orbital-angular-momentum-based oceanic turbulence links," *Appl. Opt.* **58**, 6085–6090 (2019).
40. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, 2016), pp. 770–778.
41. Y. Zhai, S. Fu, J. Zhang, X. Liu, H. Zhou, and C. Gao, "Turbulence aberration correction for vector vortex beams using deep neural networks on experimental data," *Opt. Express* **28**, 7515–7527 (2020).
42. W. Y. Wang and H.-K. Lo, "Machine learning for optimal parameter prediction in quantum key distribution," *Phys. Rev. A* **100**, 062334 (2019).
43. W. Q. Liu, P. Huang, J. Y. Peng, J. P. Fan, and G.-H. Zeng, "Integrating machine learning to achieve an automatic parameter prediction for practical continuous-variable quantum key distribution," *Phys. Rev. A* **97**, 022316 (2018).
44. P. Chaiwongkhon, K. B. Kuntz, Y. B. Zhang, A. Q. Huang, J. P. Bourgoin, S. H. Sajeed, N. Lütkenhaus, T. Jennewein, and V. Makarov, "Eavesdropper's ability to attack a free-space quantum-key-distribution receiver in atmospheric turbulence," *Phys. Rev. A* **99**, 062315 (2019).
45. Z. Q. Wang, R. Malane, and B. Burnett, "Satellite-to-Earth quantum key distribution via orbital angular momentum," *Phys. Rev. Appl.* **14**, 064031 (2020).
46. J. Liu, P. Wang, X. Zhang, Y. He, X. Zhou, H. Ye, Y. Li, S. Xu, S. Chen, and D. Fan, "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," *Opt. Express* **27**, 16671–16688 (2019).

47. L. C. Andrews, R. L. Phillips, C. Y. Hopen, and M. A. Al-Habash, "Theory of optical scintillation," *J. Opt. Soc. Am. A* **16**, 1417–1429 (1999).
48. V. Badrinarayanan, A. Kendall, and R. Cipolla, "SegNet: a deep convolutional encoder-decoder architecture for image segmentation," *IEEE Trans. Pattern. Anal.* **39**, 2481–2495 (2017).
49. F. C. Chen, "Back-propagation neural networks for nonlinear self-tuning adaptive control," *IEEE Contr. Syst. Mag.* **10**, 44–48 (1990).
50. S. Ruder, "An overview of gradient descent optimization algorithms," *arXiv:1609.04747* (2016).
51. C. Liu, K. Pang, Z. Zhao, P. Liao, R. Zhang, H. Song, Y. Cao, J. Du, L. Li, H. Song, Y. Ren, G. Xie, Y. Zhao, J. Zhao, S. M. H. Rafsanjani, A. N. Willner, J. H. Shapiro, R. W. Boyd, M. Tur, and A. E. Willner, "Single-end adaptive optics compensation for emulated turbulence in a bi-directional 10-Mbit/s per channel free-space quantum communication link using orbital-angular-momentum encoding," *Research* **2019**, 8326701 (2019).
52. D. Rosenberg, C. Peterson, J. Harrington, P. Rice, N. Dallmann, K. Tyagi, K. McCabe, S. Nam, B. Baek, R. Hadeld, R. J. Hughes, and J. E. Nordholt, "Practical long-distance quantum key distribution system using decoy levels," *New J. Phys.* **11**, 045009 (2009).
53. R. Neo, M. Goodwin, J. Zheng, J. Lawrence, S. Leon-Saval, J. Bland-Hawthorn, and G. Molina-Terriza, "Measurement and limitations of optical orbital angular momentum through corrected atmospheric turbulence," *Opt. Express* **24**, 2919–2930 (2016).
54. V. N. Mahajan, "Strehl ratio for primary aberrations in terms of their aberration variance," *J. Opt. Soc. Am.* **73**, 860–861 (1983).
55. H. K. Lo, X. Ma, and K. Chen, "Decoy-state quantum key distribution," *Phys. Rev. Lett.* **94**, 230504 (2005).
56. D. Gottesman, H. K. Lo, N. Lutkenhaus, and J. Preskill, "Security of quantum key distribution with imperfect devices," *Quant. Inf. Comp.* **4**, 325–360 (2004).
57. W. Wang, F. Xu, and H. K. Lo, "Prefixed-threshold real-time selection method in free-space quantum key distribution," *Phys. Rev. A* **97**, 032337 (2018).

# PHOTONICS Research

## Deep plug-and-play priors for spectral snapshot compressive imaging

**SIMING ZHENG,<sup>1,2,†</sup> YANG LIU,<sup>3,†</sup> ZIYI MENG,<sup>4</sup> MU QIAO,<sup>5</sup> ZHISHEN TONG,<sup>6,7</sup> XIAOYU YANG,<sup>1,2</sup> SHENSHENG HAN,<sup>6,7</sup> AND XIN YUAN<sup>8,\*</sup>**

<sup>1</sup>Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

<sup>4</sup>Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>5</sup>New Jersey Institute of Technology, Newark, New Jersey 07102, USA

<sup>6</sup>Key Laboratory for Quantum Optics and Center for Cold Atom Physics of CAS, Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China

<sup>7</sup>Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

<sup>8</sup>Nokia Bell Labs, Murray Hill, New Jersey 07974, USA

\*Corresponding author: xyuan@bell-labs.com

Received 5 October 2020; revised 20 November 2020; accepted 23 November 2020; posted 23 November 2020 (Doc. ID 411745); published 21 January 2021

We propose a plug-and-play (PnP) method that uses deep-learning-based denoisers as regularization priors for spectral snapshot compressive imaging (SCI). Our method is efficient in terms of reconstruction quality and speed trade-off, and flexible enough to be ready to use for different compressive coding mechanisms. We demonstrate the efficiency and flexibility in both simulations and five different spectral SCI systems and show that the proposed deep PnP prior could achieve state-of-the-art results with a simple plug-in based on the optimization framework. This paves the way for capturing and recovering multi- or hyperspectral information in one snapshot, which might inspire intriguing applications in remote sensing, biomedical science, and material science. Our code is available at: <https://github.com/zsm1211/PnP-CASSI>. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.411745>

### 1. INTRODUCTION

Real scenes are spectrally rich. Capturing the color, and thus the spectral information, has been a central issue since the dawn of photography. Correspondingly, many strategies have been considered. Since the advent of solid-state imaging, the color filter array and especially the red–green–blue (RGB) bayer filter have been the dominant strategy [1]. These filter arrays usually only capture red, green, and blue bands and thus limit the spectral resolution. When the number of sampled wavelengths becomes large, bandpass filters, push-room, and other strategies may be desirable. These systems usually have limited temporal resolution due to the inherent scanning procedure. Advances in photonics and 2D materials give rise to compact solutions to single-shot spectrometers at a high spectral resolution [2–5]. More recently, it has been applied for spectral imaging via combining stacking [6], optical parallelization [7], and compressive sampling [8] strategies, where the trade-off between the spatial pixel and spectral resolution still remains a challenge. Thanks to compressive sensing (CS) [9–11] and the advent of decompressive inference algorithms over the past couple of decades, there is substantial interest in hyperspectral color

filter arrays [12–14]. Such sampling strategies capture localized coded image features and are well-matched to sparsity-based inference algorithms [15–17]. With these advanced algorithms, this technique has led to single-shot imaging for hyperspectral images (HSIs), and we dub it snapshot compressive imaging (SCI) [16,18]. In this paper, we focus on the spectral SCI, which aims to measure the  $(x, y, \lambda)$  data cube.

Spectral SCI is a hardware encoder plus software decoder system, where the hardware encoder denotes the optical system, which compresses the 3D  $(x, y, \lambda)$  data cube to a snapshot measurement on the 2D detector, and the software decoder denotes the reconstruction algorithms used to recover the 3D data cube from the snapshot measurement.

The underlying principle of the spectral SCI hardware is to modulate different bands (corresponding to different wavelengths) in the spectral data cube by different weights and then integrate the light to the sensor. To perform the modulation, which should be different for different spectral bands, various techniques have been used. The pioneer work of coded aperture snapshot spectral imaging (CASSI) [12] used a fixed mask (coded aperture) and two dispersers to implement the band-wise

modulation, termed DD-CASSI; here DD means dual disperser. Following this, the single-disperser (SD) CASSI was developed [19], which achieves modulation by removing a disperser. Following CASSI, various spectral SCI systems have been built using disperser/prism and masks [20–24]. Recently, motivated by the spectral variant responses of other media, spatial light modulators [25], ground-glass-based light field modulation [26], and scatters [27] have also been employed for spectral SCI. In addition, some compact systems have also been built [28,29].

The software decoder, i.e., the reconstruction algorithm, plays a pivotal role in spectral SCI as it outputs the desired data cube. At the beginning, optimization-based algorithms developed for inverse problems such as CS were employed. Since spectral SCI is an ill-posed problem, regularizers or priors are generally used, such as the sparsity [30] and total variation [15]. Later, the patch-based methods such as dictionary learning [25,31] and Gaussian mixture models [32] were developed for the reconstruction of spectral SCI. Recently, by utilizing the nonlocal similarity in the spectral data cube, group sparsity [17] and low-rank models [16] have been developed to achieve state-of-the-art results. The main bottleneck of these high performance iterative optimization-based algorithms is the low reconstruction speed. Since the spectral data cube is usually large-scale, sometimes it needs hours to reconstruct a spectral data cube from a snapshot measurement. This precludes the real applications of spectral SCI systems.

To address the above speed issue in optimization algorithms, and inspired by the performance of deep-learning approaches for other inverse problems [33,34], convolutional neural networks (CNNs) have been used to solve the inverse problem of spectral SCI for the sake of high speed [35–39]. These networks have led to better results than their optimization counterparts, given sufficient training data and time, which usually take days or weeks. After training, the network can output the reconstruction instantaneously and thus lead to end-to-end spectral SCI sampling and reconstruction [39]. However, these networks are usually system-specific. For example, different numbers of spectral bands exist in different spectral SCI systems. Further, due to the different designs of masks, the trained CNNs cannot be used in other systems, while retraining a new network from scratch would take a long time.

Bearing the above concerns in mind, i.e., optimization-based and deep-learning-based algorithms each have their own pros and cons, it is desirable to develop a fast, flexible, and high accuracy algorithm for spectral SCI. Fortunately, the plug-and-play (PnP) framework [40,41] has been proposed for inverse problems with provable convergence [42,43]. The idea of PnP is intuitive, since the goal is to use the state-of-the-art denoiser as a simple plug-in for recovery. The rationale here is to employ recent advanced deep denoisers [44–46] in the iterative optimization algorithm to speed up the reconstruction process. Since these denoisers are pretrained with a wide range of noise levels, the PnP algorithm is very efficient and usually only tens or hundreds of iterations would provide promising results [18]. More importantly, no training is required for different tasks and thus the same denoising network can be directly used in different systems. Therefore, PnP is a good trade-off for reconstruction quality, speed, and flexibility.

However, since most existing flexible denoising networks are designed for natural images, i.e., the gray-scale or RGB images, directly using these networks into spectral SCI systems would not lead to good results. To address this issue, in this paper, we propose training a flexible denoising network for multispectral/HSIs and then apply it to the PnP framework to solve the reconstruction problem of spectral SCI.

Our proposed approach enjoys the advantages of speed, flexibility, and high accuracy. We apply the proposed method in five different real systems (three SD-CASSI systems [39,47,48], one multispectral endomicroscopy system [36], and one ghost imaging spectral system [26]) and all of them have achieved promising results. To compare with other state-of-the-art algorithms, simulations are also conducted to provide quantitative analysis. Spectral sensor design and fabrication [2,4–8] may benefit from our method by taking inspiration from the coding mechanisms and the simple plug-in for recovery.

Note that the PnP framework has been used in other inverse problems such as video CS [18], which emphasized the theoretical analysis of PnP for SCI problems in general and used an off-the-shelf denoiser (FFDNet) [46] to demonstrate its capability in video SCI. No spectral SCI results have been shown therein because spectral SCI is more challenging in terms of its various coding mechanisms and no off-the-shelf denoiser could provide a fast, flexible, and high-accuracy solution. As a matter of fact, this observation serves as the initial motivation for this paper. Towards this end, the novelty of this paper is twofold. First, we propose a CNN-based deep spectral denoising network as the spatio-spectral prior, which is flexible in terms of data size and the input noise levels. Second, we summarize the image-plane and aperture-plane coding mechanisms for spectral SCI and use the PnP method combined with our proposed deep spectral denoising prior for both simulations and five different spectral SCI systems (including image-plane and aperture-plane coding-based ones).

The paper is organized as follows. Section 2 introduces different spectral SCI systems. The proposed PnP method is derived in Section 3. Extensive results are shown in Section 4, and Section 5 concludes the entire paper.

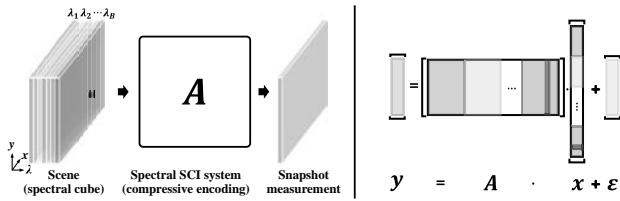
## 2. SPECTRAL SCI

The basic idea of SCI is to encode 3D or multidimensional visual information onto 2D sensor measurement. For spectral SCI, a 3D spatio-spectral data cube is encoded to form a snapshot 2D measurement on the charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) sensor, as shown in Fig. 1.

### A. SCI Forward Model

The forward model of SCI is linear. For spectral SCI, the spectral data cube of the scene  $\mathbf{X} \in \mathbb{R}^{W \times H \times B}$ , where  $W$ ,  $H$ , and  $B$  denote the width, height, and the number of spectral bands, respectively, is encoded onto a single 2D measurement  $\mathbf{Y} \in \mathbb{R}^{W \times H}$  (or similar size) via spectrally variant coding. By vectorizing the scene's spectral cube and measurement, that is,  $\mathbf{x} = \text{vec}(\mathbf{X}) \in \mathbb{R}^{WHB}$  and  $\mathbf{y} = \text{vec}(\mathbf{Y}) \in \mathbb{R}^{WH}$ , we can form a linear system for spectral SCI,

$$\mathbf{y} = \mathbf{Ax} + \boldsymbol{\varepsilon}, \quad (1)$$



**Fig. 1.** Generalized image formation (left) and the discrete matrix-form model (right) of spectral SCI. Here color denotes the corresponding spectral band.

where  $A \in \mathbb{R}^{WH \times WHB}$  and  $e \in \mathbb{R}^{WH}$  denote the sensing matrix and the measurement/sensor noise, respectively, as shown in Fig. 1.

The spatio-spectral coding mechanism is characterized by the sensing matrix (or transport matrix from the light transport perspective), i.e.,  $A$  of the optical system, where each column of the sensing matrix  $A$  is the vectorized image on the measurement plane by turning on the corresponding one voxel of the scene, as shown in the highlighted purple column of Fig. 1.

## B. Spectral SCI Systems

To encode spectral information onto a single-shot measurement, the sensing matrix must be spectrally variant. To this end, spectral SCI systems need to involve spectral dispersion devices (dispersers), like prisms, diffraction gratings, or diffusers.

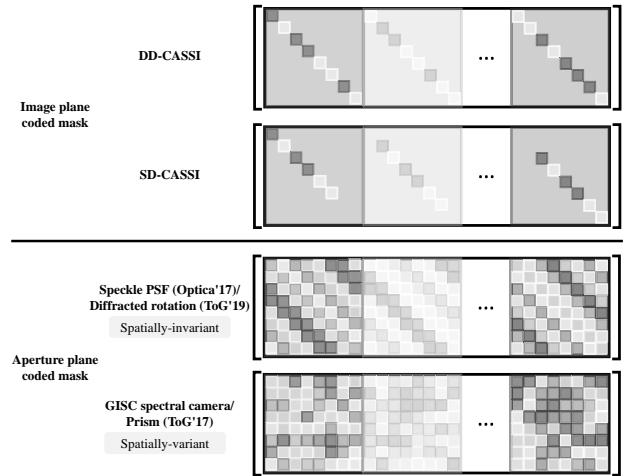
Different spectral SCI systems distinguish each other by varying the coding mechanisms, which contribute to different structures of the sensing matrices. According to the coding mechanisms, i.e., the relative position of the coded mask, spectral SCI systems could be categorized into two types, i.e., image-plane coded masks and aperture-plane coded masks. The key difference here is whether one spatio-spectral voxel (e.g., the purple voxel on the left of Fig. 1) contributes to only one element of the sensing matrix  $A$  or not.

### 1. Image-Plane Coded Mask

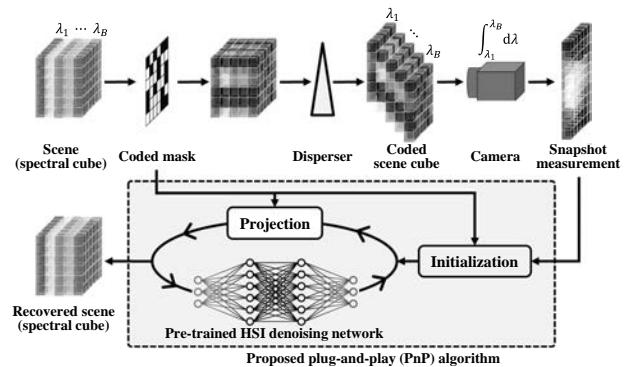
For image-plane coding, the coded mask is typically located at the conjugate image plane of the sensor plane, where one spatio-spectral voxel is directly modulated by one pixel on the coded mask and then relayed to one pixel on the detector. Therefore, there is a voxel-to-pixel mapping between the scene and the corresponding column of the sensing matrix.

As mentioned before, CASSI [12,19,47,48] was the first spectral SCI system, to the best of our knowledge. And CASSI systems can be categorized into image-plane coded masks, whether they use dual dispersers or a single disperser. The key success of CASSI is to use a coded mask for spatial coding and implement a spectral shearing with a disperser (a prism [12,19,27,47,48], a grating [20], or other spectrally variant devices like spatial light modulators (SLMs) [25,49,50]) to encode 3D spatio-spectral information onto a snapshot measurement on a 2D detector.

DD-CASSI [12] preshears the spectral cube of the scene via the first prism and then spatially encodes it using a coded mask at the image plane, where the coded spectral cube is finally unsheared to match the size of the original spectral cube via the second prism. Thereby, each voxel of the scene spectral cube would correspond to one element in the sensing matrix, and



**Fig. 2.** Comparison of image-plane coding (upper) and aperture-plane coding (lower) spectral SCI systems in terms of sensing matrix. Here each color block denotes the corresponding transport matrix at that spectral band.



**Fig. 3.** Image formation process of a typical spectral SCI system, i.e., SD-CASSI and the reconstruction process using the proposed deep PnP prior algorithm.

the encoded spectral cube is unsheared and thus has the same spatial size as the 2D measurement thanks to the usage of two complementary prisms, as shown in the first row of Fig. 2. Single disperser, or SD-CASSI [19,47] does not preshear the scene spectral cube and only performs the spatial coding and spectral shearing with a coded mask and a prism successively, as shown in the upper part of Fig. 3. In this way, the encoded spectral cube is sheared and contains some zero rows along the shearing boundaries, as shown in the second row of Fig. 2.

The common advantage of spectral SCI systems based on an image-plane coded mask is that since one spatio-spectral voxel contributes to only one element of the sensing matrix, the final sensing matrix is a concatenation of diagonal matrices, that is,

$$A = [D_1, \dots, D_B], \quad (2)$$

where  $D_b = \text{diag}[\text{vec}(C_b)] \in \mathbb{R}^{WH \times WH}$  with  $C_b$  being the (calibrated) coded mask for the  $b$ th spectral band,  $b = 1, \dots, B$ . Therefore,  $AA^\top$  is a diagonal matrix with each element the

element-wise square sum of the spectrally variant coded masks, i.e.,  $\mathbf{A}\mathbf{A}^\top = \sum_{b=1}^B \mathbf{D}_b \mathbf{D}_b^\top$ . This key property of image-plane coding-based SCI systems benefits the reconstruction algorithms significantly by reducing the computational complexity, especially for projection-based algorithms [16,51]. We will focus on the SD-CASSI case for simulations and real experiments due to the efficient hardware design.

## 2. Aperture-Plane Coded Mask

Spectral SCI systems using an aperture-plane coded mask achieve spatial encoding at the aperture plane. Each spatio-spectral voxel in the scene spectral cube is propagated to the whole sensor plane, whereas only one point is propagated for the image-plane coded mask. In this way, the sensing matrix of aperture-plane coding is a dense matrix and  $\mathbf{A}\mathbf{A}^\top$  is generally not diagonal, thus less computationally efficient for projection-based algorithms. As a general method for spectral SCI, the proposed deep PnP prior can be integrated to tackle challenges brought by various coding mechanisms (thus being flexible) by retaining efficiency at the same time; this will be discussed in Section 3.

There are two types of implementations for aperture-plane coding of a spectral SCI. The main difference is whether the point spread function (PSF) of each spatio-spectral voxel of the scene spectral cube is spatially invariant or not. Typical spatially invariant implementations are using speckles along with memory effect [52,53] and a diffractive optical element (DOE) [28] for spatially invariant PSFs, as shown in the third row of Fig. 2. Less calibration is involved for spatially invariant implementations, which would also suffer from this assumption mismatch. Spatially variant PSFs are more general, with a ghost imaging via sparsity constraints (GISC) spectral camera [26,54] and the compact prism-based spectral camera [29] as two representatives, as shown in last row of Fig. 2. We will talk about both the algorithm for aperture-coding-based spectral SCI (Section 3.A) and the experimental results on the GISC spectral camera [54] (Section 4.B.3) as well.

## 3. METHODS

Recovering 3D or multidimensional information from 2D SCI measurements is an ill-posed linear inverse problem. The main take-away from the CS [9,10,55,56] community is that sub-Nyquist sampling and reliable recovery could be achieved by constraints of the sampling/sensing matrix [55,57] and proper priors of the signal. The performance bound of the SCI-induced sensing matrix has been proved in Ref. [58]. And the fact that denoisers using deep neural networks could serve as the prior of natural images with certain constraints on the network training process is getting wide attention [43].

For the sparsity prior of the signal,  $\ell_1$  norm would be sufficient for near-optimal recovery [55,56]. For natural images, or specifically spectral images, the prior distribution of natural spectral images is needed for a good recovery. From the statistical inference perspective, we can use the maximum *a posteriori* probability (MAP) estimate, given the measurement  $\mathbf{y}$  and the forward model (likelihood function  $p_{\mathbf{y}|\mathbf{x}}$ ) to estimate the unknown signal  $\mathbf{x}$  in Eq. (1), that is,

$$\begin{aligned}\hat{\mathbf{x}} &= \arg \max_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{x}|\mathbf{y}) = \arg \max_{\mathbf{x}} \frac{p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) p_x(\mathbf{x})}{p_y(\mathbf{y})} \\ &= \arg \max_{\mathbf{x}} p_{\mathbf{y}|\mathbf{x}}(\mathbf{y}|\mathbf{x}) p_x(\mathbf{x}).\end{aligned}\quad (3)$$

Given the assumption of additive white Gaussian noise (AWGN) of the measurements  $\mathbf{e} \sim \mathcal{N}(0, \sigma_e^2)$ , the MAP form Eq. (3) can be rewritten as

$$\begin{aligned}\hat{\mathbf{x}} &= \arg \max_{\mathbf{x}} \exp \left[ -\frac{1}{2\sigma_e^2} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \log p_x(\mathbf{x}) \right] \\ &= \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_2^2 - \sigma_e^2 \log p_x(\mathbf{x}).\end{aligned}\quad (4)$$

By replacing the unknown noise variance  $\sigma_e^2$  with a noise-balancing factor  $\lambda$  and negative log prior function  $p_x(\mathbf{x})$  with a regularization term  $R(\mathbf{x})$ , Eq. (4) can be written as

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{y} - \mathbf{Ax}\|_2^2 + \lambda R(\mathbf{x}).\quad (5)$$

We further use the PnP method [40,41] based on the alternating direction method of multipliers (ADMM) [59] for image-plane coding and the two-step iterative shrinkage/thresholding (TwIST) [15] algorithm for aperture-plane coding to solve Eq. (5).

### A. PnP Method

The basic idea of PnP method for inverse problems is to use a pretrained denoiser for the desired signal as a prior. It builds on the optimization-based recovery method, where the whole inverse problem is broken into easier subproblems by handling the forward-model (data-fidelity) term and the prior term separately [59] and alternating the solutions to subproblems in an iterative manner. This is why it is called the PnP method, since the denoiser could serve as a simple plug-in for the reconstruction process. Here, for spectral SCI, we use a pre-trained HSI denoising network as the deep spectral prior and integrate it into an iterative optimization framework for reconstruction, as shown in the lower part of Fig. 3. We will start with the PnP–ADMM method for spectral SCI with image-plane coding, and then substitute the ADMM projection with TwIST for aperture-plane coding. Note that the difference lies in the “Projection” step in Fig. 3.

The proposed PnP method has guaranteed convergence for SCI with a bounded denoiser [42,43] and the assumption of estimated noise levels in a nonincreasing order [18].

#### 1. PnP–ADMM for Image-Plane Coding

The ADMM solution to the optimization problem Eq. (5) can be written as

$$\mathbf{x}^{k+1} = \arg \min_{\mathbf{x}} \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2 + \frac{\rho}{2} \|\mathbf{x} - (\mathbf{z}^k - \mathbf{u}^k)\|_2^2,\quad (6)$$

$$\mathbf{z}^{k+1} = \arg \min_{\mathbf{z}} \lambda R(\mathbf{z}) + \frac{\rho}{2} \|\mathbf{z} - (\mathbf{x}^{k+1} + \mathbf{u}^k)\|_2^2,\quad (7)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + (\mathbf{x}^{k+1} - \mathbf{z}^{k+1}),\quad (8)$$

where  $\mathbf{z}$  is an auxiliary variable,  $\mathbf{u}$  is the multiplier,  $\rho$  is a penalty factor, and  $k$  is the index of iterations. Recalling the proximal

operator [60], defined as  $\text{prox}_g(\mathbf{v}) = \arg \min_{\mathbf{x}} g(\mathbf{x}) + \frac{1}{2} \|\mathbf{x} - \mathbf{v}\|_2^2$ , the ADMM solution to SCI problem can be rewritten as

$$\mathbf{x}^{k+1} = \text{prox}_{f/\rho}(\mathbf{z}^k - \mathbf{u}^k), \quad (9)$$

$$\mathbf{z}^{k+1} = \text{prox}_{\lambda R/\rho}(\mathbf{x}^{k+1} + \mathbf{u}^k), \quad (10)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + (\mathbf{x}^{k+1} - \mathbf{z}^{k+1}), \quad (11)$$

where  $f(\mathbf{x}) = \frac{1}{2} \|\mathbf{Ax} - \mathbf{y}\|_2^2$ . Equation (9) is the Euclidean projection with a closed-form solution, i.e.,  $\mathbf{x}^{k+1} = (\mathbf{A}^\top \mathbf{A} + \rho \mathbf{I})^{-1} [\mathbf{A}^\top \mathbf{y} + \rho(\mathbf{z}^k - \mathbf{u}^k)]$ . Let  $\sigma^2 = \lambda/\rho$ , and Eq. (10) can be viewed as a denoiser  $\mathcal{D}_\sigma(\cdot)$  with  $\sigma$  as the estimated noise standard deviation.

Furthermore, recalling that  $\mathbf{AA}^\top$  is a diagonal matrix for image-plane coding,  $(\mathbf{A}^\top \mathbf{A} + \rho \mathbf{I})^{-1}$  can be calculated efficiently using the matrix inversion lemma (Woodbury matrix identity) [61], i.e.,

$$(\mathbf{A}^\top \mathbf{A} + \rho \mathbf{I})^{-1} = \rho^{-1} \mathbf{I} - \rho^{-1} \mathbf{A}^\top (\mathbf{I} + \rho \mathbf{A} \mathbf{A}^\top)^{-1} \mathbf{A} \rho^{-1}. \quad (12)$$

Then the Euclidean projection can be simplified and the final PnP–ADMM solution to the SCI problem [16,18,51] is

$$\mathbf{x}^{k+1} = (\mathbf{z}^k - \mathbf{u}^k) + \mathbf{A}^\top [\mathbf{y} - \mathbf{A}(\mathbf{z}^k - \mathbf{u}^k)] \oslash [\text{Diag}(\mathbf{A} \mathbf{A}^\top) + \rho], \quad (13)$$

$$\mathbf{z}^{k+1} = \mathcal{D}_{\hat{\sigma}_k}(\mathbf{x}^{k+1} + \mathbf{u}^k), \quad (14)$$

$$\mathbf{u}^{k+1} = \mathbf{u}^k + (\mathbf{x}^{k+1} - \mathbf{z}^{k+1}), \quad (15)$$

where  $\text{Diag}(\cdot)$  extracts the diagonal elements of the ensued matrix,  $\oslash$  denotes the element-wise division or Hadamard division, and  $\hat{\sigma}_k$  is the estimated noise standard deviation for the current ( $k$ th) iteration. Here, the noise penalty factor  $\rho$  is tuned to match the measurement (Gaussian) noise. For noiseless simulation,  $\rho$  is set to 0 or a small floating-point number. For the estimated noise standard deviation for each iteration  $\hat{\sigma}_k$ , we empirically use a large  $\hat{\sigma}_k$ , e.g., 50 out of 255 for the first several iterations (10 or 20 depending on the denoiser) and progressively shrink it during the iteration process, following Ref. [16].

For spectral SCI, we use a deep spectral denoiser as the prior, as detailed in Section 3.B. This is very straightforward for DD-CASSI. However, for SD-CASSI, there are spatial shifts between adjacent spectral bands because the spectrum is not unsheared by another disperser. Practically, we calibrate spatial shifts of all spectral bands or keep the same spatial shifts for all adjacent bands and calibrate the corresponding wavelengths. We take the spatial shifts into account by unshifting the spectral bands before applying denoising and then reshifting them back to match the forward model.

## 2. PnP–TwIST for Aperture-Plane Coding

As discussed in Section 2.B and Fig. 2, the sensing matrix of aperture-plane coding is dense and does not get  $\mathbf{AA}^\top$  a diagonal matrix. In this way, the matrix inversion lemma Eq. (12) will not help to simplify the calculation of the inverse  $(\mathbf{A}^\top \mathbf{A} + \rho \mathbf{I})^{-1}$  used in ADMM. And because of the structure

of aperture-plane coding,  $\mathbf{A}^\top \mathbf{A}$  is not well-conditioned, which makes ADMM both computationally inefficient and unstable for reconstruction.

In response to the efficiency and computation stability issues caused by ADMM projection, we use one variant of the iterative shrinkage/thresholding algorithms (ISTAs) [62], i.e., TwIST [15] for aperture-plane coding. ISTA and its variants use  $\mathbf{A}^\top$  instead of  $\mathbf{A}^\top (\mathbf{A} \mathbf{A}^\top)^{-1}$  for projection to avoid the matrix inversion of a large matrix  $\mathbf{A} \mathbf{A}^\top$ . In addition, TwIST employs another correction/acceleration step according to the conditioning of  $\mathbf{A}^\top \mathbf{A}$ , where the parameter could be tuned to match the measurement noise in real experiments. The final PnP–TwIST solution to the SCI problem is

$$\mathbf{x}^{k+1} = \mathbf{z}^k + \mathbf{A}^\top (\mathbf{y} - \mathbf{A} \mathbf{z}^k), \quad (16)$$

$$\boldsymbol{\theta}^{k+1} = \mathcal{D}_{\hat{\sigma}_k}(\mathbf{x}^{k+1}), \quad (17)$$

$$\mathbf{z}^{k+1} = (1 - \alpha) \mathbf{z}^{k-1} + (\alpha - \beta) \mathbf{z}^k + \beta \boldsymbol{\theta}^{k+1}, \quad (18)$$

where  $\alpha$  and  $\beta$  are the correction parameters depending on the eigenvalues of  $\mathbf{A}^\top \mathbf{A}$ , that is,  $\alpha = \hat{\gamma}^2 + 1$ ,  $\beta = 2\alpha/(\xi_1 + \tilde{\xi}_m)$ , whereas  $\hat{\gamma} = \frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}$ ,  $\kappa = \xi_1/\tilde{\xi}_m$ ,  $0 < \xi_1 \leq \lambda_i(\mathbf{A}^\top \mathbf{A}) \leq \xi_m$ ,  $\tilde{\xi}_m = \max\{1, \xi_m\}$ . In the experiment of GISC (Section 4.B.3), we use this PnP–TwIST due to the large scale of  $\mathbf{A}$ . After normalization of each column, we use the default setting in the TwIST code for the related parameters.

## B. Deep Spectral Denoising Prior

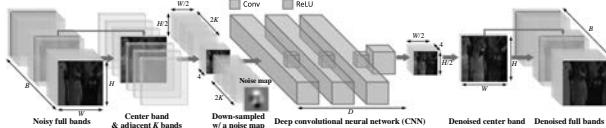
From the idea of the PnP method for linear inverse problems, we can see that a proper denoiser could serve as a prior of optimization-based approaches, where a better denoiser would contribute to higher reconstruction quality. Deep-learning-based denoisers, especially those based on CNNs for images/videos are among the state of the art. A key challenge for using deep denoisers as priors is the flexibility in terms of data size and the input noise levels. According to Eq. (14) in PnP–ADMM and Eq. (17) in PnP–TwIST, the denoiser should be adapted to different input noise levels. Inspired by the recent advance of the fast and flexible denoising CNN (FFDNet) [46] and its success applied to video SCI [18], we propose using a deep spectral denoising network as the spatio-spectral prior, that is, the deep spectral denoising prior. The network structure of the deep spectral denoising prior is shown in Fig. 4.

The spectral image denoising problem can be formulated as

$$D_\sigma(\mathbf{v}) = \text{prox}_{\sigma^2 R}(\mathbf{v}) = \arg \min_{\mathbf{x}} R(\mathbf{x}) + \frac{1}{2\sigma^2} \|\mathbf{x} - \mathbf{v}\|_2^2, \quad (19)$$

which basically learns the maximum prior probability of the HSIs, given the noisy image  $\mathbf{v}$  and the standard deviation of the Gaussian noise  $\sigma$ . Similar to the fast and flexible deep image denoiser [46,63] and the deep video denoiser [64], we perform spectral image denoising in a frame-wise manner following Ref. [65].

In order to consider the spectral correlation among adjacent bands, when denoising a center spectral frame with the size of  $W \times H$ , we take adjacent  $K$  spectral frames ( $K = 6$  in our network) as input and stack the downsampled subimages [46,63]



**Fig. 4.** Network structure of the deep spectral denoising prior.

of all  $K + 1$  frames with a noise-level map initialized as the input noise standard deviation  $\sigma$  to form a data cube of  $\frac{W}{2} \times \frac{H}{2} \times (4K + 5)$ , as shown in Fig. 4. The data cube is then transported into a CNN with 14 layers ( $D = 14$ ) of convolutional layers (Conv) and the rectified linear unit (ReLU) as the activation function (except for the last layer, where nonlinearity is not needed). We use the same size of the convolutional kernel, i.e.,  $3 \times 3$ , and zero padding to retain the image size after convolution. The number of channels for the first 13 convolutional layers is set to 128 and the last one to 4, so that the output of the CNN has a size of  $\frac{W}{2} \times \frac{H}{2} \times 4$ . This output is rearranged to arrive a single output spectral band with its original image size  $W \times H$ . Hereby, we get the denoised single-band image. After looping through all spectral bands, we can get all the spectral bands denoised. To handle the boundary case of adjacent spectral frames for the first and last few bands, we use mirror padding. Note that the key to the flexibility of our algorithm is that we need to enumerate sufficient noise levels and spectral bands during training.

### C. Training Details of Our Deep Spectral Image Denoising Network

Our denoising network is trained on the CAVE data set [66]. It contains 32 scenes with a pixel resolution of  $512 \times 512$  and 31 wavelength bands from 400 to 700 nm with a step of 10 nm. We cropped patches of size  $256 \times 256 \times 7$  from the original HSIs and employed data augmentation (rotations of  $90^\circ$ ,  $180^\circ$ ,  $270^\circ$ ; vertical flip; and combinations of the above rotation and flip operations) on the extracted patches. The total number of the patches that we finally used was 30,320. We chose seven bands during training to make sure that our denoising network could take into account the high spectral correlation between adjacent bands. We use PyTorch [67] for implementation and Adam [68] as the optimizer. The total number of training epochs is set to 500, and the batch size is set to 64 with a learning rate of  $10^{-3}$ , which decays 10 times every 100 training epochs.

Regarding the noise level  $\sigma$ , it is set to random values between 0 and 25 out of 255 during training. Training of the entire network took approximately 2 days, using a machine equipped with an Intel i5-9400F CPU, 64 GB of memory, and an Nvidia GTX 1080 Ti GPU with 11 GB RAM.

## 4. RESULTS

In this section, we verify the performance of the proposed PnP algorithm by extensive experiments. First, we conduct extensive simulations to compare PnP with other competitive methods. We then apply our PnP algorithm to data captured by real spectral SCI systems. Since different systems have different settings and parameters, the excellent results of our PnP verify the

flexibility of the proposed algorithm. Note that, for end-to-end CNN methods such as  $\lambda$ -net [37], a different network needs to be trained for each system. Moreover, since training these networks usually needs a significant amount of training data; when the system captures large-scale measurements, it will need tremendous training data and a large GPU memory, which limits the scaling performance of these end-to-end CNNs. On the other hand, our PnP algorithm can easily scale to a large data set, since the denoising is performance on patches in each iteration.

### A. Simulations

Hereby, we verify the performance of PnP by simulation using different data sets of different sizes and compare it with other popular algorithms. For the simulation data, we generate measurements following the SD-CASSI framework, as shown in the second row of Fig. 2.

#### 1. Data Sets

We employ the publicly available data sets ICVL [69] and KAIST [35] for simulation. The ICVL data are of spatial size  $1392 \times 1300$  with 31 spectral bands from 400 to 700 nm at a step of 10 nm. The KAIST data are of spatial size  $2704 \times 3376$  with 31 spectral bands from 400 to 700 nm at a step of 10 nm. We select eight scenes of each data set, shown in Fig. 5. For both data sets, we also cropped to different spatial sizes of  $256 \times 256$ ,  $512 \times 512$ , and  $1024 \times 1024$  to demonstrate the scalability of the PnP algorithm.

#### 2. Competing Methods and Comparison Metrics

We compare our proposed PnP algorithm with other popular methods, including TwIST [15], generalized alternating projection based total variation minimization (GAP-TV) [51], auto-encoder (AE) [35], and U-net [70]. Note that TwIST and GAP-TV are conventional optimization methods employing the TV prior. Though TwIST has been used for a long time for CASSI-related systems, GAP-TV has recently shown a faster convergence than TwIST. AE is a deep-learning-based algorithm that takes into account the two aspects of spectral accuracy and spatial resolution. U-net is the backbone of recently proposed deep learning for spectral compressive imaging systems including  $\lambda$ -net [37], spatial-spectral self-attention network (TSA-net) [39], and the one used in Ref. [36].

The U-net structure basically consists of two parts, the encoder part and the decoder part. Each encoder block consists of two  $3 \times 3$  convolutional layers and a  $2 \times 2$  max pooling operation. We double the feature maps during each encoder block. After four encoder blocks, we use transposed convolution operation followed by two  $3 \times 3$  convolutional layers as



**Fig. 5.** Test spectral data from (a) ICVL [69] and (b) KAIST [35] data sets used in simulation. The reference RGB images with pixel resolution  $256 \times 256$  are shown here. We crop similar regions of the whole image for spatial sizes of  $512 \times 512$  and  $1024 \times 1024$ .

one decoder block. We have doubled the feature maps during each decoder layer, too. We perform four blocks in the decoder and get the reconstructed result after a last additional  $1 \times 1$  output convolutional layer. ReLU follows each convolutional layer in both encoder and decoder as the activation function, except for the output layer, which uses the sigmoid function. Skip connections are added between the encoder blocks and decoder blocks. Similar to our denoising network, we train U-net with the CAVE data set [66]. The training process took 3 days for the spatial size of  $256 \times 256$ . Due to the long training time and GPU memory constraints, we did not train it for larger spatial sizes up to  $512 \times 512$  or  $1024 \times 1024$ . This already shows that a fixed end-to-end network such as U-net is not flexible with spatial sizes and compression ratios.

Both quantitative and qualitative metrics are used for comparison. The quantitative metrics are peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [71]. For qualitative comparison, we plot spectral frames along with spectral curves and compare them with the ground truth for visual verification. Additionally, we use Pearson correlation coefficient (corr) to assess the fidelity of recovered spectra.

### 3. Parameter Setting

From the hardware side, we use a binary random mask composed of {0,1} with the same probability. The feature size of the mask is the same as the camera. The measurement is generated following the optical path of the SD-CASSI.

For the proposed PnP algorithm, it usually needs a warm starting point to speed up the convergence. To address this, for the proposed PnP algorithm, we first run 80 iterations of GAP-TV. Since the only difference is the denoising algorithm, TV, or deep denoising, in each iteration, we only need to switch the denoising method in the flow chart, shown in Fig. 3.

The other important parameter of PnP is the noise level in each iteration. One method is to estimate the noise level in each iteration. However, this will make it computationally extensive. Therefore, similar to other PnP methods [18], we set the noise level manually in each iteration. This is also the reason we train the HSI denoising network to a wide noise range. Specifically, we set the noise level in a decreasing manner. For instance, assuming that the range of each pixel is [0,255], we set the noise level to 25 for 20 iterations, followed by 15 for 20 iterations and then tune the noise level to be smaller during the last few iterations.

### 4. Simulation Results of Different Spatial Sizes

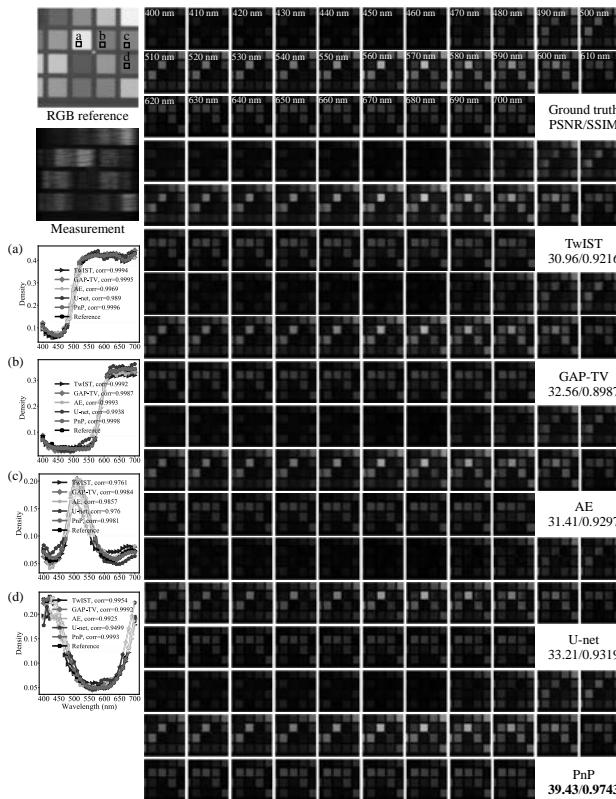
Table 1 summarizes the average results of the 16 scenes shown in Fig. 5 with different spatial sizes. It can be seen that in all these three spatial sizes, PnP always leads to the best results. In particular, PnP outperforms GAP-TV by at least 2 dB in PSNR, which is the best among other algorithms. What else stands out in the table is that AE does not perform as well as in the DD-CASSI system shown in Ref. [35]. We also tested all the above algorithms using DD-CASSI; AE can achieve better results than other algorithms except PnP.

Regarding the running time, it can be seen that for the size of  $256 \times 256$ , most methods only need about 2 min to reconstruct the spectral cube from a single measurement. At this small size, it is feasible to train a U-net for the end-to-end reconstruction. After training, the testing only needs 0.8 s,

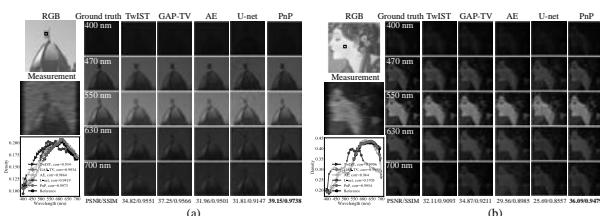
**Table 1. Average PSNR (in dB), SSIM, and Running Time (in Seconds) of 16 Simulation Scenes (8 from ICVL and 8 from KAIST) at Different Spatial Sizes Using Various Algorithms<sup>a</sup>**

Spatial Size	Data Set	TwIST			GAP-TV			AE			U-net			PnP		
		PSNR (dB)	SSIM	Running Time (s)												
256 × 256	ICVL	30.58	0.8731	156.3	32.57	0.8794	130.2	29.41	0.8711	144.2	31.13	0.8897	0.8	35.03	0.9274	132.7
	KAIST	27.32	0.8495	29.66	0.8584	26.79	0.8498	29.44	0.8941	NA	NA	NA	NA	33.21	0.9273	401.6
512 × 512	ICVL	31.82	0.8955	1380.2	33.58	0.8965	399.1	31.22	0.8969	493.6	NA	NA	NA	35.68	0.9319	401.6
	KAIST	29.09	0.8944	31.38	0.8993	29.28	0.8974	NA	NA	NA	NA	NA	NA	34.29	0.9378	401.6
1024 × 1024	ICVL	32.68	0.9159	3657.6	34.22	0.9157	1460.7	32.03	0.9158	2053.5	NA	NA	NA	36.21	0.9434	1453.6
	KAIST	31.64	0.9099	33.66	0.9134	31.05	0.9071	NA	NA	NA	NA	NA	NA	36.41	0.9433	1453.6

<sup>a</sup>NA denotes not available.



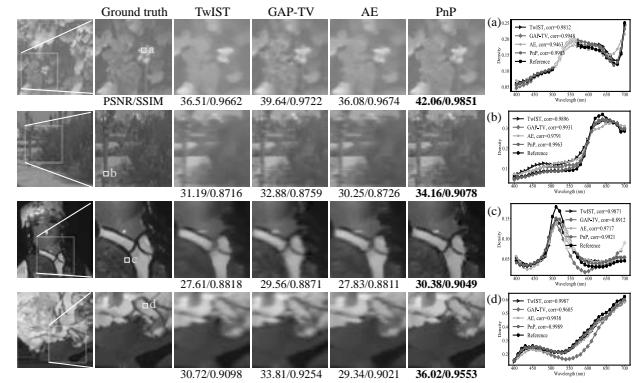
**Fig. 6.** Simulation results of color-checker with size of  $256 \times 256$  from KAIST data set compared with the ground truth. PSNR and SSIM results are also shown for each algorithm.



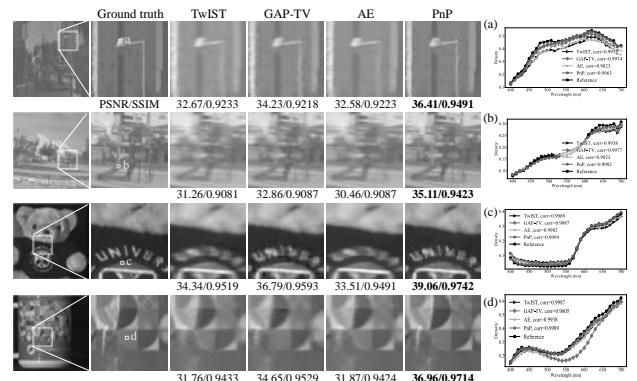
**Fig. 7.** Simulation results of exemplar scenes (top, ICVL; bottom, KAIST) with size of  $256 \times 256$  compared with the ground truth. Spectral curves of selected regions are also plotted to compare with the ground truth.

which is efficient in real applications. When the size gets larger, due to the limitation of GPU memory, we cannot train an end-to-end U-net, and thus we only show the results of the other four algorithms. It takes about 5–20 min to reconstruct a spectral cube with spatial size of  $512 \times 512$  and about 0.5 to 1 h for the size of  $1024 \times 1024$ . In summary, PnP achieves the state-of-the-art results in a relatively short time.

Figure 6 shows the results of 31 bands of each algorithm with the spatial size of  $256 \times 256$  for the scene of color-checker from KAIST data set. It can be seen clearly that PnP provides the best results. Specifically, the reconstructed frames of TwIST and GAP-TV have blocky artifacts, while the frames of AE and U-net are not clean. By contrast, the frames of PnP have fine details and sharp edges. We also plot the spectral curves of



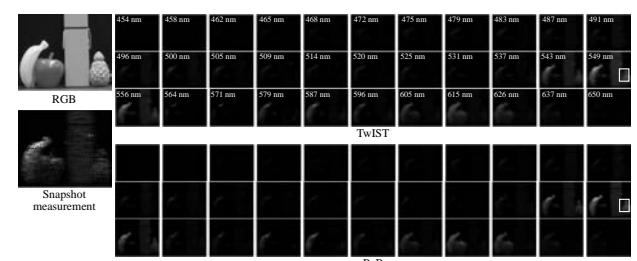
**Fig. 8.** Simulation results of four selected scenes shown in sRGB and spectral curves with spatial size of  $512 \times 512$  (shown in full size in the far left column). The spectra of the pinned (yellow) region of the close-up are shown on the right.



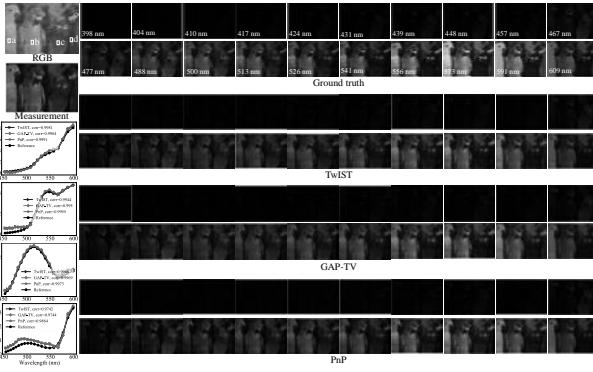
**Fig. 9.** Simulation results of four selected scenes shown in sRGB and spectral curves with spatial size of  $1024 \times 1024$  (shown in full size in the far left column). The spectra of the pinned (yellow) region of the close-up are shown on the right.

several selected regions and calculate the correlations between the reconstructed spectra and the ground truth. PnP can also provide more accurate spectra. Figure 7 plots five selected spectral frames of four other scenes. Again, it is clear that PnP provides the best results.

For other sizes of the spectral cube, in order to visualize the recovered color, we convert the spectral images to synthetic-RGB (sRGB) via the International Commission on Illumination (CIE) color-matching function [72]. The results are shown in Figs. 8 and 9, respectively, for the size of  $512 \times 512$  and  $1024 \times 1024$ .



**Fig. 10.** Real data, object SD-CASSI data ( $256 \times 210 \times 33$ ).

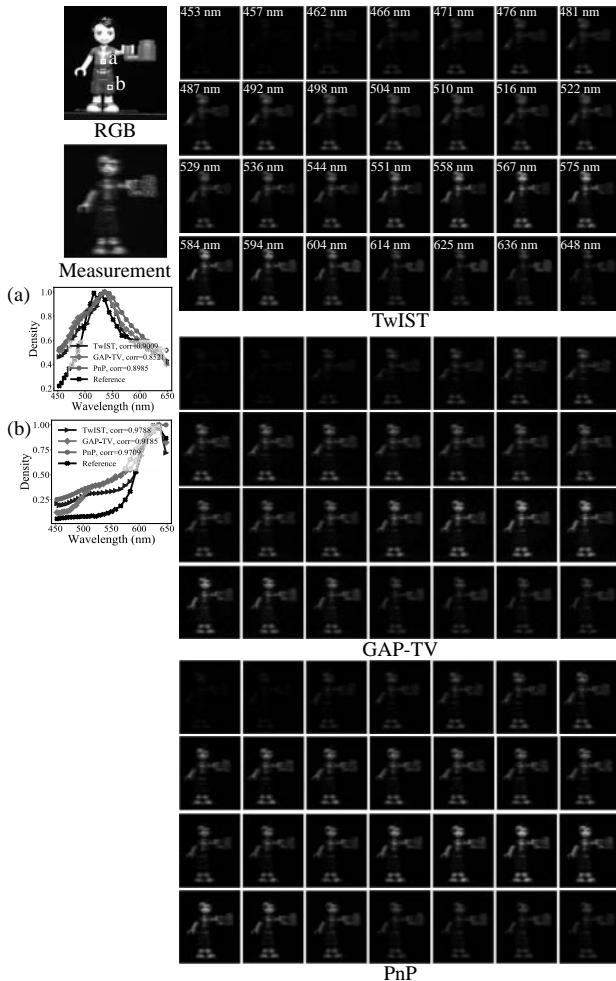


**Fig. 11.** Real data, bird SD-CASSI data ( $1021 \times 731 \times 33$ ).

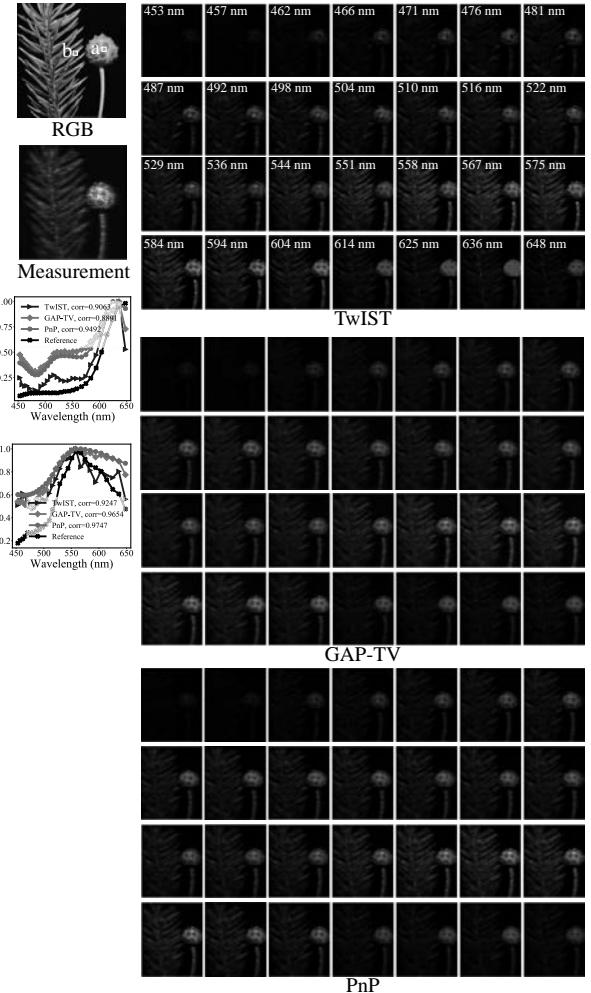
It can be observed that PnP outperforms other algorithms in both spatial details and spectral accuracy. Clear details and sharp edges can be recovered. Please refer to the zoomed regions of each scene.

## B. Real Data

In this section, we apply our proposed PnP algorithm into five real spectral SCI systems, namely, three SD-CASSI systems [39,47,48], one snapshot multispectral endomicroscopy [36],



**Fig. 12.** Real data, Lego SD-CASSI data ( $660 \times 550 \times 28$ ).



**Fig. 13.** Real data, plant SD-CASSI data ( $660 \times 550 \times 28$ ).

and a ghost spectral compressive imaging system [54]. Note that our PnP framework is using the pretrained HSI denoising network on the simulation data. Though these systems have different spatial and spectral resolutions, PnP can be used directly to all these systems. Due to the speed consideration, we only compare with TwIST and/or GAP-TV in these real data sets.

### 1. Single-Disperser CASSI

We now show three results of SD-CASSI. These measurements are captured by different systems built at different labs.

- Object data consists of 33 spectral bands, each with a size of  $256 \times 210$  pixels. The data are captured by a CASSI system built at Duke [48]. In Fig. 10, we compare the results of PnP with TwIST. We can see that fine details can be reconstructed by PnP.

- Bird data consist of 24 spectral bands, each with a size of  $1021 \times 703$  pixels, which are captured by another CASSI system built at Duke [47] along with the ground truth captured by a spectrometer. Figure 11 compares the reconstructed results of TwIST, GAP-TV, and PnP with the ground truth. We follow the similar procedure of shifting the reconstructed spectra two

bands to keep align with optical calibration, as used in Ref. [16]. For this scene, all algorithms can provide good results, but PnP achieves the clearest frames.

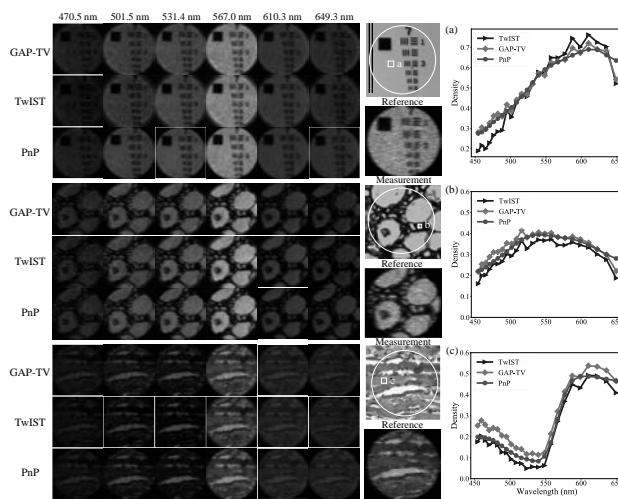
- Lego and Plant data consist of 28 spectral bands of size  $660 \times 550$ , which are captured by a recently built CASSI system at Bell Labs [39]. Figures 12 and 13 show the reconstructed results of PnP, TwIST, and GAP-TV. Clearly, PnP can provide finer details than other algorithms.

## 2. Snapshot Multispectral Endomicroscopy

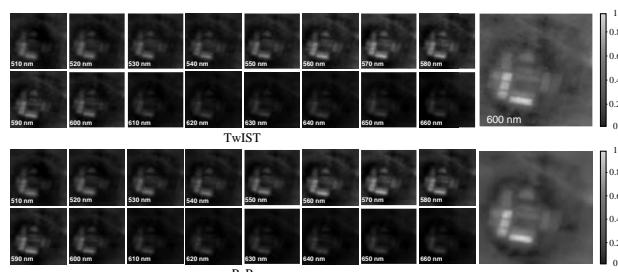
Next, we apply our PnP algorithm to the snapshot multispectral endomicroscopy system built recently [36], which is a spectral SCI system plus a fiber bundle for endoscopy. It has 24 bands in the visible bandwidth, with a spatial size of  $660 \times 660$ . We compare the results of three samples using TwIST, GAP-TV, and PnP in Fig. 14. It can be seen that both TwIST and GAP-TV lead to some noisy results, while PnP can provide clean frames.

## 3. Ghost Imaging Spectral Camera

Different from CASSI architecture, ghost imaging provides another solution to capture the spectral cube in a snapshot manner via aperture-plane coding. Hereby, we apply the PnP algorithm to the ghost imaging data captured by the system built in Ref. [54]. Since the sensing matrix of these data is large, as shown in Fig. 2, we only use the bandwidth between 510 and 660 nm with an interval of 10 nm. The spatial-spectral size of



**Fig. 14.** Real data, snapshot multispectral endomicroscopy data ( $660 \times 660 \times 24$ ).



**Fig. 15.** Real data, GISC spectral camera data ( $330 \times 330 \times 16$ ).

these data is  $330 \times 330 \times 16$ . The results of TwIST and PnP are shown in Fig. 15. It can be seen that PnP provides better results than TwIST, especially on the clean background.

## 5. CONCLUSION

We have developed a deep PnP algorithm for the reconstruction of spectral SCI. We trained a deep denoiser for hyper/multispectral images and plugged it to the ADMM and TwIST frameworks for different spectral CS systems. Importantly, a single pretrained denoiser can be applied to different systems with different settings. Therefore, our proposed algorithm is highly flexible and is ready to be used in different real applications. Extensive results on both simulation and real data captured by diverse systems have verified the performance of our proposed algorithm.

The running time scales linearly to the number of spectral bands because each spectral band is denoised individually by taking its neighboring  $K$  bands as input to the network. There are two limitations of the proposed PnP method for spectral SCI. First, it suffers from generalization issues and data set bias, as is common for supervised approaches (for example, when applying it for remote-sensing applications with hundreds of bands, fine-tuning, or retraining on the fine spectral resolution data set). Second, sometimes it needs a good initialization to start with. Since the denoiser is trained on Gaussian noise, it might have a hard time dealing with large spatial shifts in SD-CASSI. A good initialization like GAP-TV could come to the rescue. Denoisers taking the model-induced noise into account would be desirable for this PnP method.

**Disclosures.** X. Y., Nokia (E).

<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

- B. E. Bayer, "Color imaging array," U.S. patent 3,971,065 (20 July 1976).
- B. Redding, S. F. Liew, R. Sarma, and H. Cao, "Compact spectrometer based on a disordered photonic chip," *Nat. Photonics* **7**, 746–751 (2013).
- Z. Wang and Z. Yu, "Spectral analysis based on compressive sensing in nanophotonic structures," *Opt. Express* **22**, 25608–25614 (2014).
- J. Bao and M. G. Bawendi, "A colloidal quantum dot spectrometer," *Nature* **523**, 67–70 (2015).
- Z. Yang, T. Albrow-Owen, H. Cui, J. Alexander-Webber, F. Gu, X. Wang, T.-C. Wu, M. Zhuge, C. Williams, P. Wang, V. A. Zayats, W. Cai, L. Dai, S. Hofmann, M. Overend, L. Tong, Q. Yang, Z. Sun, and T. Hasan, "Single-nanowire spectrometers," *Science* **365**, 1017–1020 (2019).
- Z. Wang, S. Yi, A. Chen, M. Zhou, T. S. Luk, A. James, J. Nogan, W. Ross, G. Joe, A. Shahsafi, K. X. Wang, M. A. Kats, and Z. Yu, "Single-shot on-chip spectral sensors based on photonic crystal slabs," *Nat. Commun.* **10**, 1020 (2019).
- A. McClung, S. Samudrala, M. Torfeh, M. Mansouree, and A. Arbabi, "Snapshot spectral imaging with parallel metasystems," *Sci. Adv.* **6**, eabc7646 (2020).
- Y. Kwak, S. M. Park, Z. Ku, A. Urbas, and Y. L. Kim, "A pearl spectrometer," *Nano Lett.* (2020).
- D. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory* **52**, 1289–1306 (2006).

10. E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory* **52**, 489–509 (2006).
11. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
12. M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express* **15**, 14013–14027 (2007).
13. G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: an introduction," *IEEE Signal Process. Mag.* **31**, 105–115 (2014).
14. X. Cao, T. Yue, X. Lin, S. Lin, X. Yuan, Q. Dai, L. Carin, and D. J. Brady, "Computational snapshot multispectral cameras: toward dynamic capture of the spectral world," *IEEE Signal Process. Mag.* **33**, 95–108 (2016).
15. J. Bioucas-Dias and M. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
16. Y. Liu, X. Yuan, J. Suo, D. J. Brady, and Q. Dai, "Rank minimization for snapshot compressive imaging," *IEEE Trans. Pattern Anal. Mach. Intell.* **41**, 2990–3006 (2019).
17. L. Wang, Z. Xiong, G. Shi, F. Wu, and W. Zeng, "Adaptive nonlocal sparse representation for dual-camera compressive hyperspectral imaging," *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2104–2111 (2017).
18. X. Yuan, Y. Liu, J. Suo, and Q. Dai, "Plug-and-play algorithms for large-scale snapshot compressive imaging," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020), pp. 1447–1457.
19. A. Wagadarikar, R. John, R. Willett, and D. J. Brady, "Single disperser design for coded aperture snapshot spectral imaging," *Appl. Opt.* **47**, B44–B51 (2008).
20. X. Lin, Y. Liu, J. Wu, and Q. Dai, "Spatial-spectral encoded compressive hyperspectral imaging," *ACM Trans. Graph.* **33**, 233 (2014).
21. X. Cao, H. Du, X. Tong, Q. Dai, and S. Lin, "A prism-mask system for multispectral video acquisition," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 2423–2435 (2011).
22. H. Arguello, H. Rueda, Y. Wu, D. W. Prather, and G. R. Arce, "Higher-order computational model for coded aperture spectral imaging," *Appl. Opt.* **52**, D12–D21 (2013).
23. L. Wang, Z. Xiong, D. Gao, G. Shi, and F. Wu, "Dual-camera design for coded aperture snapshot spectral imaging," *Appl. Opt.* **54**, 848–858 (2015).
24. C. V. Correa, H. Arguello, and G. R. Arce, "Snapshot colored compressive spectral imager," *J. Opt. Soc. Am. A* **32**, 1754–1763 (2015).
25. X. Yuan, T.-H. Tsai, R. Zhu, P. Llull, D. J. Brady, and L. Carin, "Compressive hyperspectral imaging with side information," *IEEE J. Sel. Top. Signal Process.* **9**, 964–976 (2015).
26. Z. Liu, S. Tan, J. Wu, E. Li, X. Shen, and S. Han, "Spectral camera based on ghost imaging via sparsity constraints," *Sci. Rep.* **6**, 25718 (2016).
27. X. Li, J. A. Greenberg, and M. E. Gehm, "Single-shot multispectral imaging through a thin scatterer," *Optica* **6**, 864–871 (2019).
28. D. S. Jeon, S.-H. Baek, S. Yi, Q. Fu, X. Dun, W. Heidrich, and M. H. Kim, "Compact snapshot hyperspectral imaging with diffracted rotation," *ACM Trans. Graph.* **38**, 117 (2019).
29. S.-H. Baek, I. Kim, D. Gutierrez, and M. H. Kim, "Compact single-shot hyperspectral imaging using a prism," *ACM Trans. Graph.* **36**, 217 (2017).
30. M. A. T. Figueiredo, R. D. Nowak, and S. J. Wright, "Gradient projection for sparse reconstruction: application to compressed sensing and other inverse problems," *IEEE J. Sel. Top. Signal Process.* **1**, 586–597 (2007).
31. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Signal Process.* **54**, 4311–4322 (2006).
32. J. Yang, X. Liao, X. Yuan, P. Llull, D. J. Brady, G. Sapiro, and L. Carin, "Compressive sensing by learning a Gaussian mixture model from measurements," *IEEE Trans. Image Process.* **24**, 106–119 (2015).
33. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
34. X. Yuan and Y. Pu, "Parallel lensless compressive imaging via deep convolutional neural networks," *Opt. Express* **26**, 1962–1977 (2018).
35. I. Choi, D. S. Jeon, G. Nam, D. Gutierrez, and M. H. Kim, "High-quality hyperspectral reconstruction using a spectral prior," *ACM Trans. Graph.* **36**, 218 (2017).
36. Z. Meng, M. Qiao, J. Ma, Z. Yu, K. Xu, and X. Yuan, "Snapshot multispectral endomicroscopy," *Opt. Lett.* **45**, 3897–3900 (2020).
37. X. Miao, X. Yuan, Y. Pu, and V. Athitsos, " $\lambda$ -net: reconstruct hyperspectral images from a snapshot measurement," in *IEEE/CVF Conference on Computer Vision (ICCV)* (2019), pp. 4058–4068.
38. L. Wang, C. Sun, Y. Fu, M. H. Kim, and H. Huang, "Hyperspectral image reconstruction using a deep spatial-spectral prior," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 8024–8033.
39. Z. Meng, J. Ma, and X. Yuan, "End-to-end low cost compressive spectral imaging with spatial-spectral self-attention," in *European Conference on Computer Vision (ECCV)* (2020), pp. 187–204.
40. S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *IEEE Global Conference on Signal and Information Processing* (2013), pp. 945–948.
41. S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Trans. Comput. Imaging* **2**, 408–423 (2016).
42. S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: fixed-point convergence and applications," *IEEE Trans. Comput. Imaging* **3**, 84–98 (2017).
43. E. K. Ryu, J. Liu, S. Wang, X. Chen, Z. Wang, and W. Yin, "Plug-and-play methods provably converge with properly trained denoisers," *arXiv:1905.05406* (2019).
44. L. Zhang and W. Zuo, "Image restoration: from sparse and low-rank priors to deep priors," *IEEE Signal Process. Mag.* **34**, 172–179 (2017).
45. K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.* **26**, 3142–3155 (2017).
46. K. Zhang, W. Zuo, and L. Zhang, "FFDNet: toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.* **27**, 4608–4622 (2018).
47. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, "Video rate spectral imaging using a coded aperture snapshot spectral imager," *Opt. Express* **17**, 6368–6388 (2009).
48. D. Kittle, K. Choi, A. Wagadarikar, and D. J. Brady, "Multiframe image estimation for coded aperture snapshot spectral imagers," *Appl. Opt.* **49**, 6824–6833 (2010).
49. R. Zhu, T. Tsai, and D. J. Brady, "Coded aperture snapshot spectral imager based on liquid crystal spatial light modulator," in *Frontiers in Optics* (2013), paper FW1D-4.
50. T.-H. Tsai, X. Yuan, and D. J. Brady, "Spatial light modulator based color polarization imaging," *Opt. Express* **23**, 11912–11926 (2015).
51. X. Yuan, "Generalized alternating projection based total variation minimization for compressive sensing," in *IEEE International Conference on Image Processing (ICIP)* (2016), pp. 2539–2543.
52. S. K. Sahoo, D. Tang, and C. Dang, "Single-shot multispectral imaging with a monochromatic camera," *Optica* **4**, 1209–1213 (2017).
53. K. Monakhova, K. Yanny, N. Aggarwal, and L. Waller, "Spectral difusercam: lensless snapshot hyperspectral imaging with a spectral filter array," *Optica* **7**, 1298–1307 (2020).
54. J. Wu, E. Li, X. Shen, S. Yao, Z. Tong, C. Hu, Z. Liu, S. Liu, S. Tan, and S. Han, "Experimental results of the balloon-borne spectral camera based on ghost imaging via sparsity constraints," *IEEE Access* **6**, 68740–68748 (2018).
55. E. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
56. E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inf. Theory* **52**, 5406–5425 (2006).
57. E. J. Candès, "The restricted isometry property and its implications for compressed sensing," *C.R. Math.* **346**, 589–592 (2008).

58. S. Jalali and X. Yuan, "Snapshot compressed sensing: performance bounds and algorithms," *IEEE Trans. Inf. Theory* **65**, 8005–8024 (2019).
59. S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.* **3**, 1–122 (2011).
60. N. Parikh and S. Boyd, "Proximal algorithms," *Found. Trends Optim.* **1**, 127–239 (2014).
61. W. W. Hager, "Updating the inverse of a matrix," *SIAM Rev.* **31**, 221–239 (1989).
62. I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Commun. Pure Appl. Math.* **57**, 1413–1457 (2004).
63. M. Gharbi, G. Chaurasia, S. Paris, and F. Durand, "Deep joint demosaicking and denoising," *ACM Trans. Graph.* **35**, 191 (2016).
64. M. Tassano, J. Delon, and T. Veit, "FastDVDnet: towards real-time deep video denoising without flow estimation," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2020), pp. 1354–1363.
65. A. Maffei, J. M. Haut, M. E. Paoletti, J. Plaza, L. Bruzzone, and A. Plaza, "A single model CNN for hyperspectral image denoising," *IEEE Trans. Geosci. Remote Sens.* **58**, 2516–2529 (2020).
66. F. Yasuma, T. Mitsunaga, D. Iso, and S. K. Nayar, "Generalized assorted pixel camera: postcapture control of resolution, dynamic range, and spectrum," *IEEE Trans. Image Process.* **19**, 2241–2253 (2010).
67. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "PyTorch: an imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, H. Wallach, H. Larochelle, A. Beygelzimer, F. d' Alché-Buc, E. Fox, and R. Garnett, eds. (Curran Associates, 2019), pp. 8024–8035.
68. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv:1412.6980* (2014).
69. B. Arad and O. Ben-Shahar, "Sparse recovery of hyperspectral signal from natural RGB images," in *European Conference on Computer Vision* (2016), pp. 19–34.
70. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), pp. 234–241.
71. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
72. T. Smith and J. Guild, "The C.I.E. colorimetric standards and their use," *Trans. Opt. Soc.* **33**, 73–134 (1931).

# PHOTONICS Research

## High-fidelity image reconstruction for compressed ultrafast photography via an augmented-Lagrangian and deep-learning hybrid algorithm

**CHENGSHUAI YANG,<sup>1</sup> YUNHUA YAO,<sup>1,6</sup> CHENGZHI JIN,<sup>1</sup> DALONG QI,<sup>1</sup> FENGYAN CAO,<sup>1</sup> YILIN HE,<sup>1</sup> JIALI YAO,<sup>1</sup> PENGPEG DING,<sup>1</sup> LIANG GAO,<sup>2</sup> TIANQING JIA,<sup>1</sup> JINYANG LIANG,<sup>3</sup> ZHENRONG SUN,<sup>1</sup> AND SHIAN ZHANG<sup>1,4,5,7</sup>**

<sup>1</sup>*State Key Laboratory of Precision Spectroscopy, School of Physics and Electronic Science, East China Normal University, Shanghai 200062, China*

<sup>2</sup>*Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA*

<sup>3</sup>*Institut National de la Recherche Scientifique, Centre Énergie Matériaux Télécommunications, Laboratory of Applied Computational Imaging, Varennes, Québec J3X1S2, Canada*

<sup>4</sup>*Collaborative Innovation Center of Extreme Optics, Shanxi University, Taiyuan 030006, China*

<sup>5</sup>*Collaborative Innovation Center of Light Manipulations and Applications, Shandong Normal University, Jinan 250358, China*

<sup>6</sup>e-mail: yhyao@lps.ecnu.edu.cn

<sup>7</sup>e-mail: sazhang@phy.ecnu.edu.cn

Received 15 September 2020; revised 16 November 2020; accepted 2 December 2020; posted 3 December 2020 (Doc. ID 410018); published 21 January 2021

Compressed ultrafast photography (CUP) is the fastest single-shot passive ultrafast optical imaging technique, which has shown to be a powerful tool in recording self-luminous or non-repeatable ultrafast phenomena. However, the low fidelity of image reconstruction based on the conventional augmented-Lagrangian (AL) and two-step iterative shrinkage/thresholding (TwIST) algorithms greatly prevents practical applications of CUP, especially for those ultrafast phenomena that need high spatial resolution. Here, we develop a novel AL and deep-learning (DL) hybrid (i.e., AL + DL) algorithm to realize high-fidelity image reconstruction for CUP. The AL + DL algorithm not only optimizes the sparse domain and relevant iteration parameters via learning the dataset but also simplifies the mathematical architecture, so it greatly improves the image reconstruction accuracy. Our theoretical simulation and experimental results validate the superior performance of the AL + DL algorithm in image fidelity over conventional AL and TwIST algorithms, where the peak signal-to-noise ratio and structural similarity index can be increased at least by 4 dB (9 dB) and 0.1 (0.05) for a complex (simple) dynamic scene, respectively. This study can promote the applications of CUP in related fields, and it will also enable a new strategy for recovering high-dimensional signals from low-dimensional detection. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.410018>

### 1. INTRODUCTION

Ultrafast imaging has played an indispensable role in photochemistry [1,2], biomedicine [3–5], microfluidics [6], shock waves [7], and plasma physics [8]. Recently, various ultrafast imaging techniques have been developed, including compressed ultrafast photography (CUP) [9–11]. Unlike some active ultrafast imaging techniques that need specific illumination light [12–14] or a pump–probe technique that requires multiple measurements [15–17], CUP is a single-shot and passive ultrafast imaging technique. Its temporal resolution and number of frames can reach tens of femtoseconds and several hundred, respectively. Therefore, CUP has great advantages for measuring some self-luminous or non-repeatable ultrafast phenomena, which is attributed mainly to the novel model of

CUP, which combines compressed sensing (CS) theory and time–space conversion technology. So far, CUP has been successfully applied to measure light reflection and refraction [9], femtosecond temporal focusing [10], photonic Mach cones [18], dissipative solitons [19], phase-sensitive transparent objects [20], three-dimensional (3D) objects [21], ultrashort laser spatiotemporal evolution [22], and photoluminescence processes [9]. However, due to the high data compression ratio, the fidelity of reconstructed images for CUP is relatively low by the conventional two-step iterative shrinkage/thresholding (TwIST) algorithm, which limits its practicality. To improve image fidelity, a variety of methods have been proposed, such as a space- and intensity-constrained image reconstruction algorithm [23], augmented-Lagrangian (AL)-based image

reconstruction algorithm [24], plug-and-play alternating direction method of multipliers algorithm [25], optimizing the codes for CUP [26], lossless CUP [18], and multi-encoding CUP [27]. These proposed schemes can improve image fidelity to a certain extent, but there are still great challenges in measuring the complex dynamic scenes.

In image reconstruction of CUP, all selections of the sparse domain, determination of relevant iteration parameters, and denoising after iteration calculation greatly limit image fidelity. To completely solve these problems, we developed a novel image reconstruction method based on an AL and deep-learning (DL) hybrid (i.e., AL + DL) algorithm. This idea is borrowed mainly from some early algorithms, such as the AL algorithm [24,28,29], learning iteration parameters [30–33], learning sparse domain [34–37], and U-net architecture [38], but there are still many differences compared to each of the early algorithms. First, the AL + DL algorithm utilizes multiple learning transformations to seek the best sparse domain. Typically, the sparse domain in conventional TwIST and AL algorithms is determined before image reconstruction [24,39], so it is usually not optimal for one dynamic scene. In contrast, the sparse domain in the AL + DL algorithm can be optimized in multiple transformations, which is more pertinent. Second, the AL + DL algorithm takes full advantage of gradient descent (GD), DL, and AL algorithms, which simplifies the mathematical architecture to deal with the 3D tensor problem, and these advantages can reduce the cost of each iteration and decrease the number of iterations. Third, the AL + DL algorithm optimizes the relevant iteration parameters by learning the dataset, which is different from previous AL and TwIST algorithms, where these parameters are artificially predetermined. Finally, the AL + DL algorithm uses a U-net architecture containing attention layers to help denoise and retain the spatial details of the images after iteration calculation. Importantly, our theoretical simulation and experimental results show the AL + DL algorithm can obtain much higher image fidelity than conventional AL and TwIST algorithms for CUP, which strongly supports our theory.

## 2. PRINCIPLE

In CUP, a 3D dynamic scene  $I(x, y, t)$  is encoded by operator  $C$ , sheared by operator  $S$ , and integrated by operator  $T$ , and finally a two-dimensional (2D) image  $E(x', y')$  is obtained. For convenience, hereafter,  $I(x, y, t)$  is abbreviated to  $I$ , and  $E(x', y')$  is abbreviated to  $E$ . Mathematically, this process can be described as

$$E = TSCI. \quad (1)$$

For simplicity, we define  $O = TSC$ . Thus, Eq. (1) can be further written as

$$E = OI. \quad (2)$$

To recover 3D  $I$  from 2D  $E$ , we need to solve the inverse problem of Eq. (2). The number of elements in  $I$  is much larger than that in  $E$ , so the inverse problem of Eq. (2) is undetermined. The CUP strategy is to introduce a CS theory [9]. The CS theory makes full use of the sparsity of  $I$  in a certain domain to recover the original information. This sparsity in one domain means that only a few elements are nonzero, while most of the

elements are zero. Consider a case in which  $I$  has  $n$  elements and  $E$  has  $m$  elements in the original domain, and  $I$  has  $s$  non-zero elements in a sparse domain, i.e., the sparsity where  $n \gg s$  and  $n > m > s$ . Due to the fact that  $m$  is generally larger than  $s$ , this makes it possible to solve the inverse problem of Eq. (2). In a practical solution, the CS algorithm minimizes  $I$  in a sparse domain on condition of Eq. (1), which is shown as

$$\begin{cases} \min_I \Phi(I) \\ \text{s.t. } E - OI = 0 \end{cases}, \quad (3)$$

where  $\Phi(I)$  is the expression of  $I$  in the sparse domain. According to CS theory [40,41], the original dynamic scene  $I$  can be completely recovered when

$$m > f\mu^2, \quad (4)$$

where  $f$  is a constant correlated with the number of elements  $n$ ,  $\mu$  is the mutual coherence between the sparse basis of the sparse domain  $\Phi$  and operator  $O$ , determined by operators  $C$ ,  $S$ , and  $T$ . From Eq. (4), one can see that both increasing  $m$  and reducing  $s$  and  $\mu$  are feasible schemes to improve the quality of image reconstruction. However, increasing  $m$ , i.e., increasing the sampling rate, will reduce the spatial resolution or requires many streak cameras, which is impractical in the actual CUP system. Thus, reducing  $s$  and  $\mu$  is the best choice. Optimizing  $O$  can reduce only  $\mu$ , while optimizing  $\Phi$  can reduce both  $s$  and  $\mu$ ; therefore, here we employ the method of optimizing  $\Phi$ . To optimize  $\Phi$ , we impose a low-rank property on the entire dynamic scene  $I$  (tensor) with many different transformations [42], which is different from traditional methods with only one transformation. Thus, problem (3) can be further written as

$$\begin{cases} \min_{I(x,y,t)} \sum_p^q \psi_p I \\ \text{s.t. } E - OI = 0 \end{cases}, \quad (5)$$

where  $\psi_p$  represents one transformation in the sparse domain, and  $q$  denotes the total number of transformations. In transforming problem (5) from a constraint into an unconstraint, there exist two frameworks: the penalty function method and the AL method. The performance of the AL method is better than that of the penalty function method, which has been proved in previous works [24,29], and therefore, here the AL method is adopted. Thus, problem (5) can be transformed into

$$\min_I \left\{ \sum_p^q \psi_p I - \gamma(E - OI) + \frac{\zeta}{2} \|E - OI\|_F^2 \right\}, \quad (6)$$

where  $\gamma$  and  $\zeta$  are the Lagrangian multiplier and penalty parameter, respectively, which are associated with  $I$ . For convenience, problem (6) is further written as

$$\min_I \left\{ \sum_p^q \psi_p I + \frac{\zeta}{2} \left\| E - \frac{\gamma}{\zeta} - OI \right\|_F^2 \right\}. \quad (7)$$

To solve problem (7), an auxiliary variable  $J$  is introduced into problem (7), and is written as

$$\begin{cases} \min_I \left\{ \sum_p^q \psi_p J + \frac{\zeta}{2} \left\| E - \frac{\gamma}{\zeta} - OI \right\|_F^2 \right\}, \\ \text{s.t. } I = J \end{cases} \quad (8)$$

By adopting the AL method, the constrained problem (8) can be transformed into

$$\begin{aligned} \min_{I,J} & \left\{ \sum_p^q \left[ \psi_p J - \lambda_p (I - J) + \frac{\delta_p}{2} \|I - J\|_F^2 \right] \right. \\ & \left. + \frac{\zeta}{2} \left\| E - \frac{\gamma}{\zeta} - OI \right\|_F^2 \right\}, \end{aligned} \quad (9)$$

where  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_q\}$  and  $\delta = \{\delta_1, \delta_2, \dots, \delta_q\}$  are the Lagrangian multipliers and penalty parameters, respectively, which are associated with  $J$  in different transformations. By transformation, problem (9) can be further written as

$$\min_{I,J} \left\{ \sum_p^q \left( \psi_p J + \frac{\delta_p}{2} \left\| I - \frac{\lambda_p}{\delta_p} - J \right\|_F^2 \right) + \frac{\zeta}{2} \left\| E - \frac{\gamma}{\zeta} - OI \right\|_F^2 \right\}. \quad (10)$$

Problem (10) can be solved by an alternating direction method of multipliers (ADMM) based on an iteration of solving the  $I$ -subproblem and  $J$ -subproblem alternatively. However, in the  $J$ -subproblem, the sparse domains in different transformations lead to different solutions at the beginning of the iteration. Therefore, some independent auxiliary variables  $\mathbf{W} = \{W_1, W_2, \dots, W_q\}$  are introduced for each transformation, and thus problem (10) can be written as

$$\min_{I,W} \left\{ \sum_p^q \left( \psi_p W_p + \frac{\delta_p}{2} \left\| I - \frac{\lambda_p}{\delta_p} - W_p \right\|_F^2 \right) + \frac{\zeta}{2} \left\| E - \frac{\gamma}{\zeta} - OI \right\|_F^2 \right\}. \quad (11)$$

To solve problem (11), the ADMM is also adopted to solve the  $I$ -subproblem and  $\mathbf{W}$ -subproblem alternatively. In the  $k$ th iteration, the  $I$ -subproblem can be written as

$$\begin{aligned} I^k = \arg \min_I & \left\{ \sum_p^q \frac{\delta_p^k}{2} \left\| I - \frac{\lambda_p^k}{\delta_p^k} - W_p^{k-1} \right\|_F^2 \right. \\ & \left. + \frac{\zeta^k}{2} \left\| E - \frac{\gamma^k}{\zeta^k} - OI \right\|_F^2 \right\}, \end{aligned} \quad (12)$$

and the  $\mathbf{W}$ -subproblem can be written as

$$W_p^k = \arg \min_{W_p} \psi_p W_p + \frac{\delta_p^k}{2} \left\| I^k - \frac{\lambda_p^k}{\delta_p^k} - W_p \right\|_F^2. \quad (13)$$

The  $I$ -subproblem in Eq. (12) is a quadratic regularized least-squares problem, and its direct solution is given in a closed form as

$$\begin{cases} I^k = \text{HARD} \left[ \zeta^k O^T E - \gamma^k O^T + \left( \sum_p^f \delta_p^k W_p^{k-1} - \lambda_p^k \right) \right], \\ \text{HARD} = \left( \zeta^k O^T O + \sum_p^f \delta_p^k I_{\text{idem}} \right)^{-1}, \end{cases} \quad (14)$$

where  $I_{\text{idem}}$  is an identity operator. It is expensive to compute **HARD** due to large data, so here the GD algorithm is used to solve the  $I$ -subproblem. The main shortcoming of the GD algorithm is that the number of iterations is very large because it is difficult to choose the step size. Some methods have been developed to obtain a better step size through much computation, such as the Barzilai and Borwein (BB) method [43,44]. Here, a learning method is used to seek the optimized step size. In our method, the number of iterations is much less than that of the BB method. Based on the GD algorithm, the solution to Eq. (12) can be expressed as

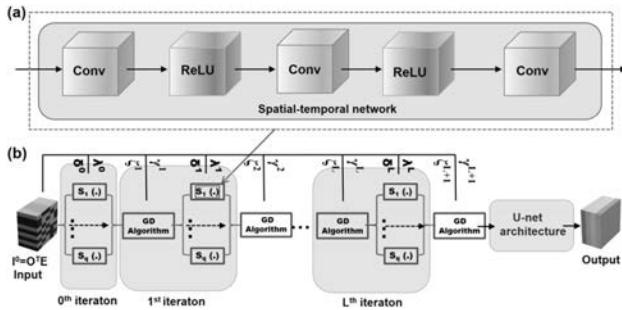
$$\begin{aligned} I^k = I^{k-1} - \alpha^k & \left[ \sum_p^q \delta_p^k \left( I^{k-1} - \frac{\lambda_p^k}{\delta_p^k} - W_p^{k-1} \right) \right. \\ & \left. - \zeta^k O^T \left( E - \frac{\gamma^k}{\zeta^k} - OI^{k-1} \right) \right], \end{aligned} \quad (15)$$

where  $\alpha$  is the step size in the GD algorithm, and here the GD algorithm is utilized to calculate Eq. (15). From Eqs. (15) and (13), one can see that the solution to problem (6) depends mainly on  $W_p$  instead of  $I$ . For convenience, the solution to  $W_p$  can be further written as

$$W_p^k = S_p \left( I^k - \frac{\lambda_p^k}{\delta_p^k} \right). \quad (16)$$

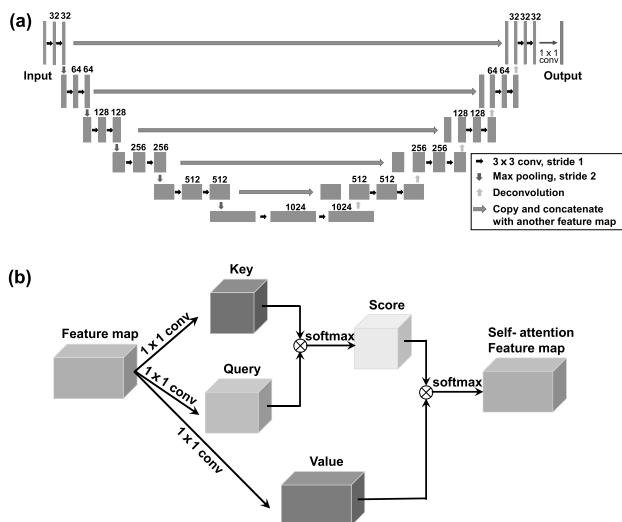
To obtain the solver  $S_p(\cdot)$ , the traditional algorithms usually employ the explicit handmade image prior as the sparse domain, such as a total variation (TV) prior and a wavelet prior [29,36]. However, the hand-crafted image prior has no pertinence for one dynamic scene, so it is not the best sparse domain. Here, we propose to learn the solver  $S_p(\cdot)$  by convolutional neural networks. The architecture of learning solver  $S_p(\cdot)$  is a spatial-temporal network, which is utilized to exploit the sparse domain from spatial and temporal correlation. This network consists of two sets of convolutional layers followed by a rectified linear unit (ReLU) layer and a single convolutional layer, as shown in Fig. 1(a), which is motivated by a recent work on image spatial super-resolution [45].

The general framework of the AL + DL algorithm is shown in Fig. 1(b). Compared with the conventional TwIST or AL algorithm, we optimize the sparse domain and some relevant iteration parameters  $\{\alpha^k, \delta^k, \zeta^k\}$  by end-to-end training. The optimized sparse domain by specific training can greatly reduce the sparsity  $s$  and coherence  $\mu$ , which is very helpful for high-fidelity image reconstruction, as shown in Eq. (4). To help denoise and retain more details after the iteration, we add U-net architecture containing self-attention, as shown in Fig. 2. The U-net has five times downsampling and upsampling, as shown in Fig. 2(a). In particular, we have two times convolution operations with stride 1 after downsampling or upsampling. Also, we impose self-attention to the layer that has



**Fig. 1.** Data flow chart of the AL + DL algorithm. (a) Solver  $S_p(\cdot)$  in a sparse domain; (b) general framework by connecting each iteration in a sequence order. Here, each  $S_i(\cdot)$  is calculated in parallel to producing  $W_i$ , and GD algorithm is employed to calculate  $I$ .

128 feature maps before deconvolution, which can help the architecture learn the long-range similarity easily, as shown in Fig. 2(b). Here, U-net allows the network to propagate the context information to some higher resolution layers, which has been successfully utilized to recover 3D information from 2D information in the spectral images [38]. Meanwhile, the self-attention mechanism, which has been recently proposed in computer vision tasks [46–49], can be used to exploit both the non-local similarity of spatial textures and the long-range temporal similarity, because the self-attention can help networks focus on some specific details and form some local specific feature. By embedding U-net architecture, the mean peak signal-to-noise ratio (PSNR) value of all the images in our simulating dynamic scenes increases by 0.81 dB, while the mean structural similarity index (SSIM) value increases by 0.007. Therefore, the AL + DL algorithm can retain more spatial details and finally achieve higher image fidelity than conventional AL and TwIST algorithms in theory. To facilitate researchers in citing and using our AL + DL algorithm, the codes are available at <https://github.com/integritynoble/ALDL-algorithm>.

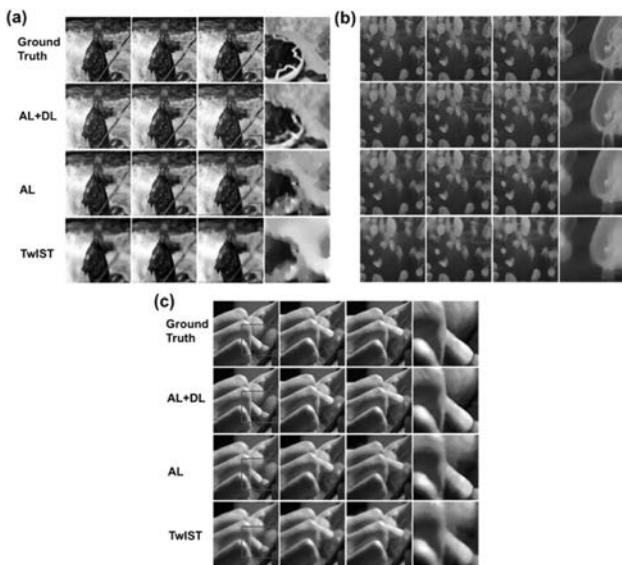


**Fig. 2.** (a) U-net architecture in the AL + DL algorithm; (b) self-attention model.

### 3. THEORETICAL SIMULATIONS

To validate the superior performance of the AL + DL algorithm in CUP, we perform three theoretical simulations and two experiments. In image reconstruction, TensorFlow is employed to implement the AL + DL algorithm on an NVIDIA GeForce GTX 2080Ti GPU with 11 GB device memory. Initially, the size of all images should be resized to  $32N \times 32M$  due to five ( $2^5 = 32$ ) times downsampling and upsampling in the U-net architecture, but the number  $K$  of frames is not limited, which indicates that the dynamic scene should have  $32N \times 32M \times K$  cube, where  $N$ ,  $M$ , and  $K$  can be adjusted according to the real dynamic scene. In fact, the resizing of the image has no side effect on the dynamic scene, because the size of images can be set to be larger than the actual one by padding zeros. When learning the model, the relevant iteration parameters are set as follows: all initial elements in Lagrangian multipliers  $\gamma$  and  $\lambda = \{\lambda_1, \lambda_2, \dots, \lambda_q\}$  are set to zero, initial  $I$  is set as  $O^T E$ , the number of iterations is 11, maximum running epoch is 280, and the initial learning rate is 0.008. Meanwhile, a rooted square-mean-error (RMSE) is used as the training loss, which is minimized by the Adam optimizer [50]. In each iteration, the values of Lagrangian multipliers  $\gamma$  and  $\lambda$  are calculated with the AL algorithm [24,28,29,51]. In our theoretical simulations, we chose three kinds of dynamic scenes with different complexities to test the ability of the AL + DL algorithm in the image reconstruction of CUP, and each dynamic scene contains eight frames. The three dynamic scenes are boatman [52], ocean animal [53], and finger [54]. Here, the boatman scene has some droplets and subtle textures, and therefore the relevant images are difficult to compress, representing the complex scene, while the finger scene contains only finger movement; thus, the relevant images are easy to compress, representing a simple scene. Usually, the inverse of the lossless compression ratio of images can be used to illustrate the complexity of a dynamic scene [55]. For each dynamic scene, 512 relevant pictures are utilized to train the model, and  $k$ -fold cross-validation is used to track the training effect. Here, this set of pictures is divided into two parts: one is used as training images, and the other is used as test images. To train the model, the 512 pictures are grouped and then combined into many small videos, and each video contains eight pictures, which corresponds to the frame number of each dynamic scene. Here, only one picture is replaced in each video compared to the previous video. Also, these original videos are randomly partitioned into eight equal-sized sub-videos in the eight-fold cross-validation. To show the superiority of the AL + DL algorithm, the AL and TwIST algorithms are also used for reconstruction based on the TV domain, which are used mostly for CUP [9–11,18–24, 56,57]. The reconstructed images of the boatman, ocean animal, and finger by the AL + DL, AL, and TwIST algorithms are shown in Fig. 3, together with the ground truth for comparison.

Here, only three representative pictures are selected, and an interesting area in each dynamic scene is enlarged for observation. Spatial details in the boatman, ocean animal, and finger can be clearly observed by the AL + DL algorithm, while these details are submerged by the AL and TwIST

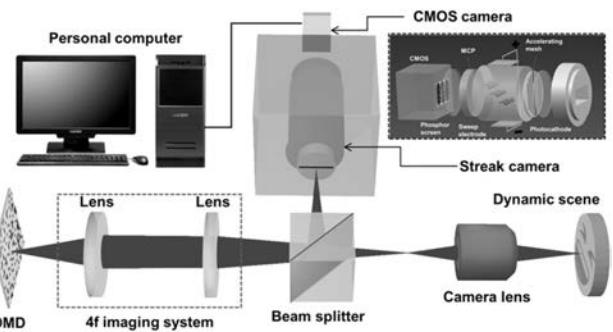


**Fig. 3.** Reconstructed results of (a) boatman, (b) ocean animal, and (c) finger by the AL + DL (second row), AL (third row), and TwIST (fourth row) algorithms, together with the ground truth (first row) for comparison. The last column is the enlarged image in the corresponding red squares.

algorithms, which is disadvantageous for high-spatial-resolution imaging of a dynamic scene. To intuitively compare the improved efficiency in image fidelity by the AL + DL algorithm, we calculate PSNR and SSIM, and the calculated results are given in Table 1. Compared to the AL and TwIST algorithms, both PSNR and SSIM by the AL + DL algorithm are significantly improved. Here, PSNR (SSIM) is increased by at least 4.35 dB (0.136) for the boatman, 5.47 dB (0.114) for the ocean animal, and 9.78 dB (0.051) for the finger. Based on these calculated results, a rule can be found, which is, the simpler the spatial structure of the dynamic scene, the higher the improvement efficiency of PSNR, while the improvement efficiency of SSIM shows the opposite behavior. This phenomenon should be related to the sparsity of the dynamic scene; the simpler dynamic scene usually has higher sparsity, and vice versa. PSNR is based on a logarithmic function, which is not very well matched to perceived visual quality, but SSIM is based on visible structures in the image. Thus, PSNR has high improvement efficiency for a simple dynamic scene (i.e., finger), while SSIM has high improvement efficiency for a complex dynamic scene (i.e., boatman). In addition, the AL + DL algorithm can reconstruct a dynamic scene in only a few seconds,

**Table 1. Average PSNR (in dB) and SSIM by Different Image Reconstruction Algorithms in Different Dynamic Scenes**

Scene	AL + DL		AL		TwIST	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Boatman	28.50	0.836	24.15	0.700	22.47	0.589
Ocean animal	30.47	0.916	25.00	0.802	24.72	0.781
Finger	42.00	0.983	32.22	0.932	28.56	0.894



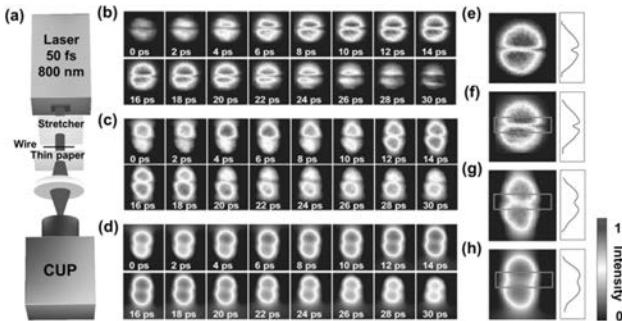
**Fig. 4.** System configuration of CUP. DMD, digital micromirror device; CMOS, complementary metal–oxide–semiconductor.

which is much shorter than the AL and TwIST algorithms, which need tens of seconds; the computing efficiency is improved by an order of magnitude, which is very beneficial in practical applications of CUP.

#### 4. EXPERIMENTAL RESULTS

Besides the above theoretical simulations, we also experimentally verify the superiority of the AL + DL algorithm on image reconstruction of CUP. The system configuration of CUP is given in Fig. 4. The dynamic scene is imaged via a camera lens and a 4f imaging system. On the image plane, a digital micro-mirror device (DMD) (Texas Instruments, DLP LightCrafter) is used to encode the dynamic scene in the spatial domain with a pseudo-random binary pattern, as encoding operator  $C$ . Through the collection of the same 4f imaging system and the reflection of a beam splitter, the encoded dynamic scene is vertically deflected by a streak camera (Hamamatsu, C7700), as shearing operator  $S$ . Finally, a complementary metal–oxide–semiconductor (CMOS) camera (Hamamatsu, ORCA-flash4.0) is employed to detect the encoded and deflected dynamic scene, as integrating operator  $T$ . Combining the measured image by CMOS and the codes on DMD, the original dynamic scene is reconstructed by the AL + DL, AL, and TwIST algorithms. For the training data of the AL + DL algorithm, we simulated the dynamic scenes based on the static images recorded without operators  $C$  and  $S$ .

First, we measure the temporal evolution of a spatially modulated picosecond laser spot, and the experimental design is shown in Fig. 5(a) [24]. The output 50 fs (full width at half maximum, FWHM) laser pulse from a Ti:sapphire amplifier is broadened to about 16 ps by a stretcher, and a thin wire is used to divide the laser spot into two components in space to obtain such a dynamic scene with special spatial structure. The spatially modulated laser spot illuminates a thin white paper, and a small fraction of photons can pass through the thin white paper. Thus, the temporal evolution behavior of a spatially modulated laser spot can be measured by our CUP system with a frame rate of 500 billion frames per second (fps). In this dynamic scene, the signal strength changes, while the spatial structure remains unchanged. The reconstructed images by the AL + DL, AL, and TwIST algorithms are shown in Figs. 5(b)–5(d), respectively. Compared with the AL and

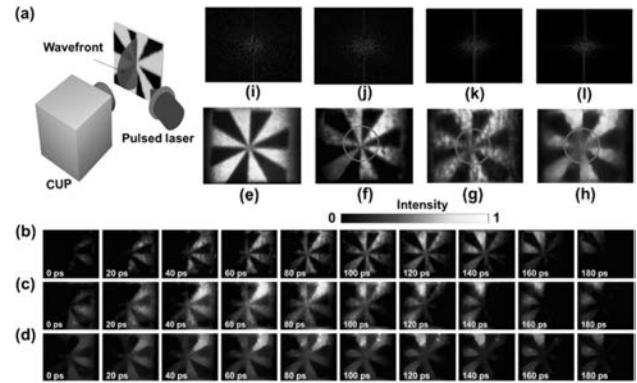


**Fig. 5.** Measuring temporal evolution of a spatially modulated picosecond laser spot. (a) Experimental design. (b)–(d) Reconstructed results by the AL + DL, AL, and TwIST algorithms, respectively. (e) Measured static image by external CCD. (f)–(h) Extracted images from (b)–(d), respectively, at the time of 14 ps; curves on the right are the integration results of the corresponding images along the horizontal direction.

TwIST algorithms, the reconstructed images by the AL + DL algorithm have a clearer spatial shape and less background noise. To further compare the image fidelity by the three algorithms, we chose the reconstructed images at a time of 14 ps to compare with the static image, as shown in Figs. 5(e)–5(h).

Here, the static image is achieved by external CCD measurement without encoding operator  $C$  and shearing operator  $S$ , as shown in Fig. 5(e). Meanwhile, the intensities of Figs. 5(e)–5(h) are also integrated along the horizontal direction, and the calculated results are given on the right of the relative images. The AL + DL algorithm can retain very high image fidelity, but the AL and TwIST algorithms cause a certain degree of image distortion. The fundamental reason should be the mismatch of the sparse domain in image reconstruction. More importantly, like the static image, the blocked part in the laser spot (see light blue squares) can be clearly distinguished by the AL + DL algorithm, where an obvious valley in the intensity curve is observed, but not by either the AL or TwIST algorithm, especially the TwIST algorithm.

In the first experiment in Fig. 5(a), the spatial shape of the dynamic scene remains unchanged. In the second experiment, we measure the wavefront movement by obliquely illuminating a collimated femtosecond laser pulse on a transverse fan pattern, where both the signal strength and spatial shape in the dynamic scene change. The experimental design is presented in Fig. 6(a). A 7 ps (FWHM) laser pulse after collimation obliquely illuminates a transverse fan pattern with an angle of  $\sim 30^\circ$  to the surface normal. Our CUP system faces the pattern surface and collects the scattered photons from the pattern scene. Here, the shearing velocity of the streak camera is 0.66 km/s; thus, the imaging speed is 50 billion fps, i.e., 20 ps exposure time in theory [9]. The reconstruction images by the AL + DL, AL, and TwIST algorithms are presented in Figs. 6(b)–6(d), respectively. As expected, the spatial shape of the fan can be displayed in the whole process of wavefront movement by the AL + DL algorithm for image reconstruction, while it is blurred by the AL and TwIST algorithms due to the artifacts in the image reconstruction. To better evaluate the image reconstruction effect of the three



**Fig. 6.** Measuring wavefront movement by obliquely illuminating a collimated femtosecond laser pulse on a transverse fan pattern. (a) Experimental design. (b)–(d) Reconstructed results by the AL + DL, AL, and TwIST algorithms, respectively. (e) Measured static image by external CCD. (f)–(h) Integrated images from (b)–(d), respectively. (i)–(l) Results of Fourier transform from (e)–(h), respectively.

algorithms, the reconstructed images in Figs. 6(b)–6(d) are integrated and compared to the static image measured by an external CCD, as shown in Figs. 6(e)–6(h). Similar to the static image, the whole outline of the fan in the integrated image via the AL + DL algorithm is clear, but it is a little fuzzy by the AL and TwIST algorithms, especially for the center part of the fan (green circles). To intuitively illustrate the spatial resolution, the images in Figs. 6(e)–6(h) are processed via Fourier transform, and the calculated results are shown in Figs. 6(i)–6(l). As can be seen, the AL + DL algorithm can obtain high-frequency information, which is almost the same as the static image, while the high-frequency information is lost for the AL and TwIST algorithms. In general, high-frequency information represents the fine structure in the spatial domain. Therefore, compared to the AL and TwIST algorithms, the AL + DL algorithm has great advantages in observing the spatial details of a complex dynamic scene.

## 5. DISCUSSION

The AL + DL algorithm is a data-driven method, which can optimize the sparse domain and relevant iteration parameters by learning instead of hand-crafted determination. For CS, the sparse domain is the core part that determines the sparsity and affects mainly the coherence. Thus, the sparse domain almost determines the image reconstruction quality. In general, the learning method can seek better sparse domain and iteration parameters, and therefore the AL + DL algorithm can get higher image fidelity than conventional AL and TwIST algorithms. Because of learning the sparse domain and iteration parameters, the AL + DL algorithm has high robustness and allows the encoding operator  $C$  to be different in training and testing processes, while the pure neural network algorithms cannot, such as deep fully connected networks [58], ReconNet [59], DR<sup>2</sup>-Net [60],  $\lambda$ -net [38], and DeepCubeNet [61]. Also, the AL + DL algorithm embeds a GD algorithm into tensor

computation, which involves massive data. In calculation, the GD algorithm does not easily find the appropriate step size, so it needs to perform many iterations, i.e., the convergence speed is low. To decrease the number of iterations, data scientists prefer Newton's method or a conjugate gradient algorithm by increasing the cost of each iteration [62]. However, some mathematicians seek a better step size in the GD algorithm to decrease the number of iterations by increasing the cost of each iteration, such as the BB method. Here, we utilize the GD algorithm to calculate the large data by a data-driven method based on the learning model, which can find the optimal step size to decrease the number of iterations without increasing the cost of each iteration and make the gradient show better orthogonality. It is noted that the AL + DL algorithm needs just 15 iterations, while the corresponding traditional algorithm based on the BB method needs more than 100 iterations.

As shown in Figs. 3, 5, and 6, compared to the AL and TwIST algorithms, the AL + DL algorithm shows great advantages in image reconstruction accuracy, but it also inherits the shortcoming of the data-driven method, i.e., the dependence on a learning dataset. In image reconstruction, these images in the dataset should have some similarities to those in the dynamic scene. An inappropriate training dataset may lead to results worse than those obtained by the AL and TwIST algorithms. In some special dynamic scenes, it may be difficult to find a similar dataset for training. In this case, it is feasible to increase the sampling rate  $m$ , such as lossless-CUP or multi-encoding CUP. Moreover, it is also a good idea to optimize the codes, which is similar to optimizing the sparse domain, which can reduce coherence. However, the AL + DL algorithm cannot be adopted directly to optimize the codes, because here the codes are considered as constant. Optimizing the codes demands that the mathematical architecture regard the codes as a variable; thus, the whole AL + DL architecture needs to be redesigned. In the future, we will strive to seek some new algorithms to simultaneously optimize the codes, sparse domain, and iteration parameters by learning the dataset.

## 6. CONCLUSION

In summary, we have developed a new AL + DL algorithm to realize high-fidelity image reconstruction for CUP. In our method, there are four key points: (1) optimizing the sparse domain in multiple transformation; (2) optimizing the relevant calculation parameters in the iteration process; (3) employing the GD algorithm to improve computing efficiency; (4) embedding the U-net architecture to help denoise. Key points (1), (2), and (4) are implemented by the DL method, and improving key point (3) also needs the DL method. However, the whole framework is determined by the AL method, which combines these four key points. Thus, the AL + DL algorithm not only utilizes the training neural networks, but also has some potential mathematical interpretations. More importantly, these results from theoretical simulations and experimental measurements show that the AL + DL algorithm is superior to conventional AL and TwIST algorithms in image fidelity and computing efficiency. Additionally, the AL + DL algorithm is a

simple mathematical architecture, so it is easy to extend to other high-dimensional tensor fields. In future studies, we will continue to search for better image reconstruction algorithms for CUP to achieve super-high image fidelity.

**Funding.** National Natural Science Foundation of China (11727810, 11774094, 11804097, 91850202); Science and Technology Commission of Shanghai Municipality (19560710300, 20ZR1417100).

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- P. R. Poulin and K. A. Nelson, "Irreversible organic crystalline chemistry monitored in real time," *Science* **313**, 1756–1760 (2006).
- P. Hockett, C. Z. Bisgaard, O. J. Clarkin, and A. Stolow, "Time-resolved imaging of purely valence-electron dynamics during a chemical reaction," *Nat. Phys.* **7**, 612–615 (2011).
- R. Horstmeyer, H. Ruan, and C. Yang, "Guidestar-assisted waveform-shaping methods for focusing light into biological tissue," *Nat. Photonics* **9**, 563–571 (2015).
- J. W. Borst and A. J. Visser, "Fluorescence lifetime imaging microscopy in life sciences," *Meas. Sci. Technol.* **21**, 102002 (2010).
- H. R. Petty, "Spatiotemporal chemical dynamics in living cells: from information trafficking to cell physiology," *Biosystems* **83**, 217–224 (2006).
- T. M. Squires and S. R. Quake, "Microfluidics: fluid physics at the nanoliter scale," *Rev. Mod. Phys.* **77**, 977–1026 (2005).
- N. Šiaulys, L. Gallais, and A. Melninkaitis, "Direct holographic imaging of ultrafast laser damage process in thin films," *Opt. Lett.* **39**, 2164–2167 (2014).
- R. L. Kodama, P. A. Norreys, K. Mima, A. E. Dangor, R. G. Evans, H. Fujita, Y. Kitagawa, K. Krushelnick, T. Miyakoshi, and N. Miyanaga, "Fast heating of ultrahigh-density plasma as a step towards laser fusion ignition," *Nature* **412**, 798–802 (2001).
- L. Gao, J. Liang, C. Li, and L. V. Wang, "Single-shot compressed ultrafast photography at one hundred billion frames per second," *Nature* **516**, 74–77 (2014).
- J. Liang, L. Zhu, and L. V. Wang, "Single-shot real-time femtosecond imaging of temporal focusing," *Light Sci. Appl.* **7**, 42 (2018).
- D. Qi, S. Zhang, C. Yang, Y. He, F. Cao, J. Yao, P. Ding, L. Gao, T. Jia, and J. Liang, "Single-shot compressed ultrafast photography: a review," *Adv. Photon.* **2**, 014003 (2020).
- K. Nakagawa, A. Iwasaki, Y. Oishi, R. Horisaki, A. Tsukamoto, A. Nakamura, K. Hirosawa, H. Liao, T. Ushida, and K. Goda, "Sequentially timed all-optical mapping photography (STAMP)," *Nat. Photonics* **8**, 695–700 (2014).
- T. Suzuki, R. Hida, Y. Yamaguchi, K. Nakagawa, T. Saiki, and F. Kannari, "Single-shot 25-frame burst imaging of ultrafast phase transition of Ge<sub>2</sub>Sb<sub>2</sub>Te<sub>5</sub> with a sub-picosecond resolution," *Appl. Phys. Express* **10**, 092502 (2017).
- Y. Lu, T. T. Wong, F. Chen, and L. Wang, "Compressed ultrafast spectral-temporal photography," *Phys. Rev. Lett.* **122**, 193904 (2019).
- A. Velten, T. Willwacher, O. Gupta, A. Veeraraghavan, M. G. Bawendi, and R. Raskar, "Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging," *Nat. Commun.* **3**, 745 (2012).
- A. H. Zewail, "Four-dimensional electron microscopy," *Science* **328**, 187–193 (2010).
- A. Barty, S. Boutet, M. J. Bogan, S. Hau-Riege, S. Marchesini, K. Sokolowski-Tinten, N. Stojanovic, H. Ehrke, A. Cavalleri, and S. Düsterer, "Ultrafast single-shot diffraction imaging of nanoscale dynamics," *Nat. Photonics* **2**, 415–419 (2008).
- J. Liang, C. Ma, L. Zhu, Y. Chen, L. Gao, and L. V. Wang, "Single-shot real-time video recording of a photonic Mach cone induced by a scattered light pulse," *Sci. Adv.* **3**, e1601814 (2017).

19. J. C. Jing, X. Wei, and L. V. Wang, "Spatio-temporal-spectral imaging of non-repeatable dissipative soliton dynamics," *Nat. Commun.* **11**, 2059 (2020).
20. T. Kim, J. Liang, L. Zhu, and L. V. Wang, "Picosecond-resolution phase-sensitive imaging of transparent objects in a single shot," *Sci. Adv.* **6**, e6200 (2020).
21. J. Liang, L. Gao, P. Hai, C. Li, and L. V. Wang, "Encrypted three-dimensional dynamic imaging using snapshot time-of-flight compressed ultrafast photography," *Sci. Rep.* **5**, 15504 (2015).
22. F. Cao, C. Yang, D. Qi, J. Yao, Y. He, X. Wang, W. Wen, J. Tian, T. Jia, and Z. Sun, "Single-shot spatiotemporal intensity measurement of picosecond laser pulses with compressed ultrafast photography," *Opt. Laser Eng.* **116**, 89–93 (2019).
23. L. Zhu, Y. Chen, J. Liang, Q. Xu, L. Gao, C. Ma, and L. V. Wang, "Space-and intensity-constrained reconstruction for compressed ultrafast photography," *Optica* **3**, 694–697 (2016).
24. C. Yang, D. Qi, F. Cao, Y. He, X. Wang, W. Wen, J. Tian, T. Jia, Z. Sun, and S. Zhang, "Improving the image reconstruction quality of compressed ultrafast photography via an augmented Lagrangian algorithm," *J. Opt.* **21**, 035703 (2019).
25. Y. Lai, Y. Xue, C. Y. Côté, X. Liu, A. Laramée, N. Jaouen, F. Légaré, L. Tian, and J. Liang, "Single-shot ultraviolet compressed ultrafast photography," *Laser Photon. Rev.* **14**, 2000122 (2020).
26. C. Yang, D. Qi, X. Wang, F. Cao, Y. He, W. Wen, T. Jia, J. Tian, Z. Sun, and L. Gao, "Optimizing codes for compressed ultrafast photography by the genetic algorithm," *Optica* **5**, 147–151 (2018).
27. C. Yang, D. Qi, J. Liang, X. Wang, F. Cao, Y. He, X. Ouyang, B. Zhu, W. Wen, and T. Jia, "Compressed ultrafast photography by multi-encoding imaging," *Laser Phys. Lett.* **15**, 116202 (2018).
28. C. Li, "An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing," Master dissertation (Rice University, 2010).
29. M. V. Afonso, J. M. Bioucas-Dias, and M. A. Figueiredo, "An augmented Lagrangian approach to the constrained optimization formulation of imaging inverse problems," *IEEE Trans. Image Process.* **20**, 681–695 (2010).
30. Y. Yang, J. Sun, H. Li, and Z. Xu, "ADMM-CSNet: a deep learning approach for image compressive sensing," *IEEE Trans. Pattern Anal.* **42**, 521–538 (2018).
31. J. Zhang and B. Ghanem, "ISTA-Net: interpretable optimization-inspired deep network for image compressive sensing," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 1828–1837.
32. J. Ma, X. Liu, Z. Shou, and X. Yuan, "Deep tensor ADMM-net for snapshot compressive imaging," in *Proceedings of the IEEE International Conference on Computer Vision* (2019), pp. 10223–10232.
33. K. Monakhova, J. Yurtsever, G. Kuo, N. Antipa, K. Yanny, and L. Waller, "Learned reconstructions for practical mask-based lensless imaging," *Opt. Express* **27**, 28075–28090 (2019).
34. Q. Xie, Q. Zhao, D. Meng, and Z. Xu, "Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery," *IEEE Trans. Pattern Anal. Mach. Intell.* **40**, 1888–1902 (2017).
35. Y. Wang, J. Peng, Q. Zhao, Y. Leung, X. Zhao, and D. Meng, "Hyperspectral image restoration via total variation regularized low-rank tensor decomposition," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing* **11**, 1227–1243 (2017).
36. L. Wang, C. Sun, Y. Fu, M. H. Kim, and H. Huang, "Hyperspectral image reconstruction using a deep spatial-spectral prior," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 8032–8041.
37. Z. Wu, Y. Sun, A. Matlock, J. Liu, L. Tian, and U. S. Kamilov, "SIMBA: scalable inversion in optical tomography using deep denoising priors," *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sensing* **14**, 1163–1175 (2020).
38. X. Miao, X. Yuan, Y. Pu, and V. Athitsos, "Lambda-net: reconstruct hyperspectral images from a snapshot measurement," in *IEEE/CVF International Conference on Computer Vision (ICCV)* (IEEE, 2019), pp. 4058–4068.
39. J. M. Bioucas-Dias and M. A. Figueiredo, "A new TwIST: two-step iterative shrinkage/thresholding algorithms for image restoration," *IEEE Trans. Image Process.* **16**, 2992–3004 (2007).
40. E. J. Candès and T. Tao, "Near-optimal signal recovery from random projections: universal encoding strategies?" *IEEE Trans. Inform. Theory* **52**, 5406–5425 (2006).
41. E. J. Candès, J. K. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Commun. Pure Appl. Math.* **59**, 1207–1223 (2006).
42. X. Liu and X. Wang, "Fourth-order tensors with multidimensional discrete transforms," arXiv:1705.01576 (2017).
43. J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA J. Numer. Anal.* **8**, 141–148 (1988).
44. M. Raydan, "Convergence properties of the Barzilai and Borwein gradient method," Ph.D. dissertation (Rice University, 1991).
45. B. Lim, S. Son, H. Kim, S. Nah, and K. Mu Lee, "Enhanced deep residual networks for single image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 136–144.
46. L. Yue, X. Miao, P. Wang, B. Zhang, X. Zhen, and X. Cao, "Attentional alignment networks," in *29th British Machine Vision Conference* (2018), pp. 1–14.
47. S. Min, X. Chen, Z. Zha, F. Wu, and Y. Zhang, "A two-stream mutual attention network for semi-supervised biomedical segmentation with noisy labels," in *Proceedings of the AAAI Conference on Artificial Intelligence* (2019), pp. 4578–4585.
48. Y. Li, Z. Xiao, X. Zhen, and X. Cao, "Attentional information fusion networks for cross-scene power line detection," *IEEE Geosci. Remote Sens. Lett.* **16**, 1635–1639 (2019).
49. Y. Huang, X. Cao, X. Zhen, and J. Han, "Attentive temporal pyramid network for dynamic scene classification," in *Proceedings of the AAAI Conference on Artificial Intelligence* (2019), pp. 8497–8504.
50. D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).
51. S. H. Chan, R. Khoshabeh, K. B. Gibson, P. E. Gill, and T. Q. Nguyen, "An augmented Lagrangian method for video restoration," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2011), pp. 941–944.
52. [https://drive.google.com/file/d/1PGav\\_1Yju2lfXIXQsxnkvMPu\\_j\\_EWm0B/view](https://drive.google.com/file/d/1PGav_1Yju2lfXIXQsxnkvMPu_j_EWm0B/view).
53. <https://www.pexels.com/video/blue-jellyfish-856882/>.
54. <https://www.pexels.com/video/a-person-using-a-smart-cellular-phone-3143531/>.
55. H. Yu and S. Winkler, "Image complexity and spatial information," in *Fifth International Workshop on Quality of Multimedia Experience (QoMEX)* (IEEE, 2013), pp. 12–17.
56. C. Yang, D. Qi, F. Cao, Y. He, J. Yao, P. Ding, X. Ouyang, Y. Yu, T. Jia, and S. Xu, "Single-shot receive-only ultrafast electro-optical deflection imaging," *Phys. Rev. Appl.* **13**, 024001 (2020).
57. C. Yang, F. Cao, D. Qi, Y. He, P. Ding, J. Yao, T. Jia, Z. Sun, and S. Zhang, "Hyperspectrally compressed ultrafast photography," *Phys. Rev. Lett.* **124**, 023902 (2020).
58. M. Iliadis, L. Spinoulas, and A. K. Katsaggelos, "Deep fully-connected networks for video compressive sensing," *Digit. Signal Process.* **72**, 9–18 (2018).
59. K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "ReconNet: non-iterative reconstruction of images from compressively sensed measurements," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 449–458.
60. H. Yao, F. Dai, S. Zhang, Y. Zhang, Q. Tian, and C. Xu, "DR<sup>2</sup>-Net: deep residual reconstruction network for image compressive sensing," *Neurocomputing* **359**, 483–493 (2019).
61. D. Gedalin, Y. Olnikine, and A. Stern, "DeepCubeNet: reconstruction of spectrally compressive sensed hyperspectral images with deep neural networks," *Opt. Express* **27**, 35811–35822 (2019).
62. J. Nocedal and S. Wright, *Numerical Optimization* (Springer, 2006).

# PHOTONICS Research

## Smart ring resonator-based sensor for multicomponent chemical analysis via machine learning

ZHENYU LI,<sup>1,2,†</sup> HUI ZHANG,<sup>2,†</sup> BINH THI THANH NGUYEN,<sup>2</sup> SHAOBO LUO,<sup>2</sup> PATRICIA YANG LIU,<sup>2</sup> JUN ZOU,<sup>2</sup> YUZHI SHI,<sup>2</sup> HONG CAI,<sup>3</sup> ZHENCHUAN YANG,<sup>1</sup> YUFENG JIN,<sup>1</sup> YILONG HAO,<sup>1,5</sup> YI ZHANG,<sup>2,4,6</sup> AND AI-QUN LIU<sup>2,7</sup> 

<sup>1</sup>National Key Laboratory of Science and Technology on Micro/Nano Fabrication, Institute of Microelectronics, Peking University, Beijing 100871, China

<sup>2</sup>Quantum Science and Engineering Centre, Nanyang Technological University, Singapore 639798, Singapore

<sup>3</sup>Institute of Microelectronics, A\*STAR (Agency for Science, Technology and Research), Singapore 138634, Singapore

<sup>4</sup>School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore

<sup>5</sup>e-mail: haoyl@pku.edu.cn

<sup>6</sup>e-mail: yi\_zhang@ntu.edu.sg

<sup>7</sup>e-mail: eaqliu@ntu.edu.sg

Received 6 October 2020; revised 29 October 2020; accepted 18 November 2020; posted 20 November 2020 (Doc. ID 411825); published 21 January 2021

We demonstrate a smart sensor for label-free multicomponent chemical analysis using a single label-free ring resonator to acquire the entire resonant spectrum of the mixture and a neural network model to predict the composition for multicomponent analysis. The smart sensor shows a high prediction accuracy with a low root-mean-squared error ranging only from 0.13 to 2.28 mg/mL. The predicted concentrations of each component in the testing dataset almost all fall within the 95% prediction bands. With its simple label-free detection strategy and high accuracy, the smart sensor promises great potential for multicomponent analysis applications in many fields. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.411825>

### 1. INTRODUCTION

Ring resonators are in an emerging class of versatile and highly sensitive photonic sensors that use recirculating light confined within a microcavity to detect the changes in surrounding biological, physical, and chemical environments [1–3]. They are well suited for integrated sensing systems because of their high sensitivity, compact size, label-free detection, real-time monitoring capability, low sample consumption, multiplexing capability, and resistance to electromagnetic interference [4,5]. Ring resonators could be easily fabricated into a large array for multiplexed detection. In theory, each individually addressable ring resonator in the array could be labeled with a unique capture agent to detect a specific target of interest in a mixture [6]. The capture agent labeled on the ring resonator ensures the sensing specificity. However, this approach presents a practical challenge—the ring resonators would be so closely spaced in the array that it would be impractical to selectively label each ring resonator while avoiding contaminating the adjacent ones without employing expensive and bulky microdispensers [7]. Therefore, a simple unlabeled multiplexed detection strategy is highly desired for ring resonator-based sensors. Indeed,

the ring resonator is well recognized for its capability of label-free sensing. When target molecules come into proximity of the ring resonator, the resonant peak shifts to a different wavelength, and the degree of this resonant shift reflects the target concentration. This label-free sensing strategy works well for single-component analysis but shows poor performance for multicomponent analysis due to its lack of specificity. With the endpoint measurement of the resonant peak shift as the only parameter for sensing, the ring resonator is incapable of distinguishing a specific component from a mixture. One conventional solution to this problem is to extract the component of interest from the mixture and measure all components one at a time. However, the extraction processes are often tedious, laborious, and time-consuming [8]. Moreover, not all components can be extracted from a mixture. A more popular approach of multicomponent analysis is based on multivariate calibration of spectral information obtained by using UV spectrophotometry, Raman spectroscopy, nuclear magnetic resonance spectroscopy, and other kinds of spectroscopy [9–11]. This approach requires a large database containing the spectra of each individual component in their pure forms. To identify all components, the

superimposed spectrum of the mixture is decomposed into the spectra of individual components using algorithms such as multiple linear regression, principle component regression, and partial least squares [12,13]. One major assumption made by this approach is that the spectrum of the mixture is a linear superposition of the spectra of individual components, which is not always substantiated. This approach also has difficulty in identifying components with substantial overlapping spectra and indistinguishable features. Although several models have been proposed to distinguish those spectra by introducing additional handcrafted features, they have not been widely adopted in practice [14,15].

Instead of using features handcrafted by users, machine learning identifies characteristic features from the dataset on its own. It is able to build data-driven models with its unique learning ability. In recent years, machine learning, especially deep learning [16], has received a lot of attention and has redefined data science [17,18]. The rapid development of machine learning has enabled a wide range of data-driven sensing applications. For example, research in drug discovery and development employs machine learning to explore the causal association between drugs, biomarkers, and diseases [19,20], which promotes data-driven decision-making and has the potential to speed up drug development and reduce failure rates. Machine learning is also used to predict drug–drug interactions during clinical trials, which reduces adverse drug reactions and healthcare costs [21]. Machine learning has also become a crucial technique for device and material development [22–25]. Researchers have attempted to optimize the design of metamaterial on demand with machine learning [26–28]. Various types of sensors, such as microring resonators and surface plasmon resonance-based sensors [29,30], have benefited from machine learning. Machine learning is employed to boost the selectivity of gas sensors [31] and improve the performance of low-cost and mobile plasmonic sensing platforms by reducing the inter-device variability [32].

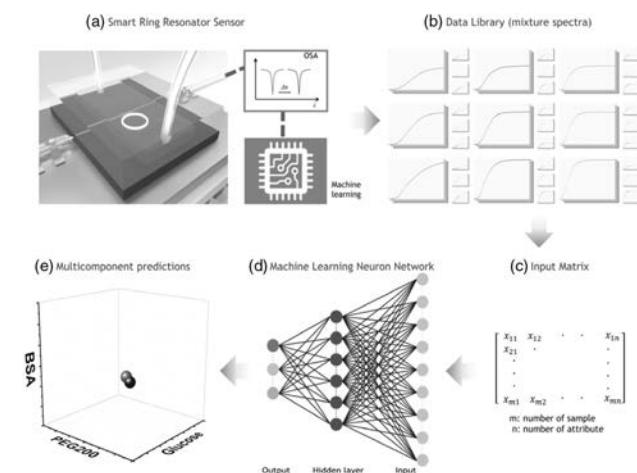
In this paper, we demonstrate a machine learning–enabled smart sensor based on a ring resonator for multicomponent chemical analysis. The smart sensor integrates a ring resonator and microfluidic network onto a single chip using silicon photonic fabrication technology. Instead of using an array of ring resonators selectively labeled with capture agents for multiplexing, a single label-free ring resonator is used to sense all components of interest in a mixture by analyzing the resonant spectrum of the mixture using a machine learning model. The machine learning model based on artificial neural network is trained to decompose the resonant spectrum of the mixture into spectra of the individual components. Instead of solely relying on the resonant peak shift, the machine learning algorithm extracts features from the entire spectrum for target identification and quantification. As a proof of concept, we have accurately quantified glucose, polyethylene glycol (PEG) 200, and bovine serum albumin (BSA) mixed at arbitrary ratios using this smart sensor with a low root-mean-squared error (RMSE) ranging only from 0.13 to 2.28 mg/mL over the tested range of 0–30 mg/mL. The reported smart sensor significantly simplifies the ring resonator–based sensing platform for multicomponent chemical analysis and shows a high quantification

accuracy. Although this work just presents a simple model system for proof of concept, it promises great potential for multicomponent analysis of complex samples in micro- and nanodevices.

## 2. METHODS AND MATERIALS

### A. Sensing System and Data Analysis

The ring resonator–based smart sensor [Fig. 1(a)] consisted of a silicon ring resonator chip (details are shown in Appendix A) that was integrated with a microfluidic network, a 1550 nm light source, a polarization controller, an optical spectrum analyzer (OSA), and a data acquisition (DAQ) circuit, all packaged into a portable system. The light source was tuned to TE polarization via a fiber polarization controller and coupled into the bus waveguide through a lensed fiber. After propagating through the chip, the light that exited from the through port of the sensing ring was coupled out of the bus waveguide via the other lensed fiber to the OSA for spectrum acquisition. The gap between the bus waveguide and the ring ought to be kept in the range of hundreds of nanometers, so that the light propagating along the waveguide could be coupled into the adjacent ring resonator to form a whispering-gallery mode for resonance detection under critical conditions. A piece of oxygen plasma-treated polydimethylsiloxane (PDMS) film was bonded to the ring resonator chip to form a microfluidic channel, the depth of which was 1  $\mu\text{m}$  and the width was 100  $\mu\text{m}$  [33–36]. The resonator chip was first rinsed with isopropyl alcohol (IPA), ethanol, and deionized (DI) water, followed by oxygen plasma cleaning (PDC-002, Harrick Plasma) for 5 min to remove possible surface contamination. The chip was passivated to prevent molecule adsorption by being immersed in undiluted Sigmacote (SL2, Sigma) solution for 5 min with shaking immediately after oxygen plasma cleaning. After the incubation, the chip was dried at room temperature in a sterilized hood for one hour and washed with sterilized



**Fig. 1.** Schematic illustration of smart sensor framework. (a) Acquisition of resonant spectrum using ring resonator–based smart sensor. (b) Resonant spectra are collected into a data library. (c) Transformation of resonant spectra to matrices. (d) Neural network training with transformed resonant spectra. (e) Composition prediction with neural network model.

double-distilled water to remove hydrochloric acid (HCl) byproduct. The chip was dried at 90°C in an oven for 30 min before use.

The data library contained resonant spectra of mixtures comprising PEG200, glucose, and BSA mixed at various concentrations [Fig. 1(b)]. To measure the resonant spectra, the running buffer (Ancillary Reagent Kit 2, DuoSet) was first injected into the microfluidic channel at 30  $\mu\text{L}/\text{min}$  for 2 min. The sample was then injected, and the resonant wavelength of the ring resonator was monitored over time to obtain the resonant spectrum. A total of 343 mixtures were measured with the concentration of each component varying from 0 to 30 mg/mL with an increment of 5 mg/mL, while each composition was measured at 5 different flow rates ranging from 10 to 90  $\mu\text{L}/\text{min}$  with an increment of 20  $\mu\text{L}/\text{min}$ . The microfluidic channel was washed with ample running buffer between measurements.

### B. Training of Neural Network for Composition Detection

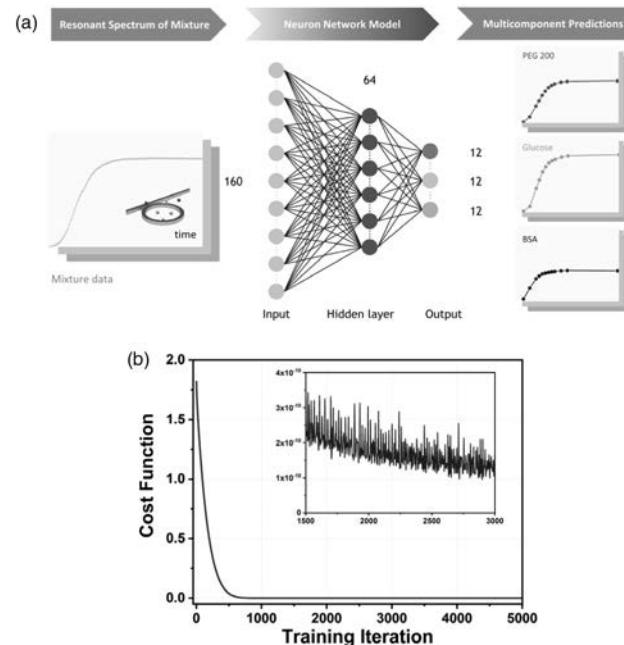
The data library was divided into training and testing datasets, and each dataset was divided into five subgroups based on the flow rate. Each spectrum in the datasets was labeled with 36 attributes that encoded the concentrations of the three components—PEG200, glucose, and BSA—in the mixture, with 12 attributes for each. The 12 attributes for each component at a particular concentration were extracted from the resonant spectrum of the component in its pure form. The resonant spectra of the mixtures were transformed into matrices [Fig. 1(c)] as the input of the neural network [Fig. 1(d)]. The neural network was trained to decompose the resonant spectrum of the mixture into resonant spectra of individual components and predict the respective concentration of each component [Fig. 1(e)].

A three-layer perceptron neural network (one input layer, one output layer, and a single hidden layer) was adopted to implement a regressor to predict the concentration of each component in a mixture [Fig. 2(a)]. The input layer had 160 neurons, the hidden layer had 64 neurons, and the output layer had 36 neurons. The collected datasets included 343 samples, of which 229 samples were used for training, and the remaining 114 samples were used for the blind testing of the neural network model. Each sample was measured at five different flow rates, and therefore the total number of training samples was  $343 \times 5$ . The training and testing datasets were randomly selected from the spectrum library without overlap, which guaranteed that the trained model did not get any information about the testing dataset. The input datasets had a dimension of 160 attributes, and the output had a dimension of 36 attributes, with 12 attributes for each component. The hidden layer had 64 neurons, and thus the matrix sizes for the input layer, hidden layer, and output layer were  $160 \times 64$ ,  $64 \times 64$ , and  $64 \times 36$ , respectively. The three layers were fully connected and activated by a rectified linear units (ReLU) function. The training model was set up using TensorFlow [37] and trained over a GPU server with one Nvidia GeForce RTX 2080 card and one Intel Xeon CPU E5-2650. The cost function was defined by the mean square error (MSE), and the optimizer was set to the RMSPropOptimizer with a learning rate of 0.0005

and a decay rate of 0.9. The learning rate and decay rate were selected by trial and error. The cost function was plotted against the number of iterations [Fig. 2(b)]. It quickly descended to a value close to zero after  $\sim 700$  iterations, which indicated the low error of the model on the training set, and no overfitting was observed up to 5000 iterations. Our successful blind test on the  $114 \times 5$  testing samples was unknown to the trained model. All evidence proved that the number of training samples was sufficient. Different from other machine learning applications on published datasets, we used a high-quality dataset experimentally collected. The  $343 \times 5$  samples and their corresponding labels were of high quality and reliable resolution. With this high-quality dataset, a relatively small number of training samples were able to generate satisfactory machine learning models for accurate prediction. The RMSE is a frequently used way to measure the error of a model in predicting quantitative data. The deviation between prediction and ground truth was evaluated over  $n_{\text{samples}}$  by the RMSE, which was expressed as

$$\text{RMSE}(y, \hat{y}) = \sqrt{\frac{1}{n_{\text{samples}}} \sum_{i=0}^{n_{\text{samples}}-1} (y_i - \hat{y}_i)^2}, \quad (1)$$

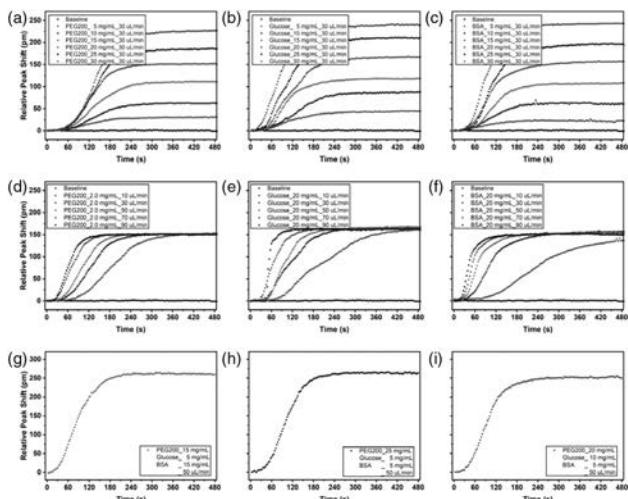
where  $\hat{y}_i$  was the predicted value of the  $i$ th sample, and  $y_i$  was the corresponding true value.



**Fig. 2.** Neural network of the smart sensor. (a) The resonant spectrum of an unknown sample mixture is acquired with the ring resonator and fed into the trained neural network. The number of neurons in the input, hidden, and output layers is 160, 64, 36, respectively. The 36 attributes of the output are divided into three groups, each with 12 attributes. Each group of 12 attributes is used to predict the concentration of one component in the mixture. The composition of the unknown sample is determined by the predicted concentrations of all three components. (b) Cost function versus training interactions for samples at 10  $\mu\text{L}/\text{min}$  as an example.

### 3. RESULTS AND DISCUSSIONS

The light is coupled into the ring resonator through fiber alignment [38,39]. The light at resonant wavelength is strictly confined in the ring resonator, which results in periodic resonant peaks in a spectrum measured at the through port of the bus waveguide. The cross section of the waveguide in our device is 450 nm by 220 nm, which could confine single-mode light very well. On the other hand, the evanescent field is able to leak from the waveguide to interact with matter that alters the effective refractive index surrounding the ring resonator. This effect is reflected as a shift in the resonant wavelength in the measured spectrum. The ring resonator is an exceptional optical sensor because of its higher  $Q$  value, high sensitivity, ease of fabrication, multiplexity capability, and label-free detection [40]. During operation, the ring resonator is immersed in the running buffer that flows stably in the microfluidic channel, and a stable resonant peak is observed. As the sample flows over the surface of the ring resonator, the resonant peak would shift as a result of the difference in refractive index between the sample solution and the running buffer. This principle is used to analyze chemicals dissolved in the solution. The proposed smart sensor combines the advantage of the highly sensitive ring resonator and neural network to empower intelligent label-free multicomponent chemical analysis. The representative resonant spectra of individual components under a constant flow rate of 30  $\mu\text{L}/\text{min}$  are shown in Figs. 3(a)–3(c). The concentrations of PEG200 [Fig. 3(a)], glucose [Fig. 3(b)], and BSA

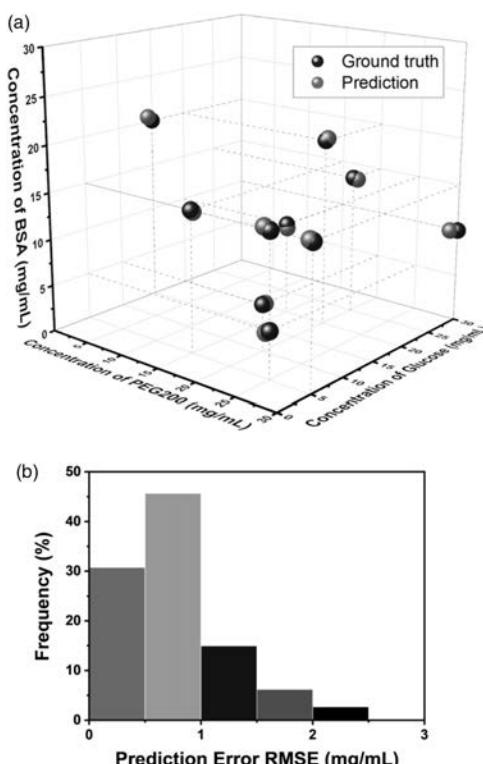


**Fig. 3.** Resonant spectra measured by the ring resonator under various conditions. (a)–(c) Resonator spectra for PEG200, glucose, and BSA solution under a constant flow rate of 30  $\mu\text{L}/\text{min}$ . The concentration ranges from 0 to 30 mg/mL with an increment of 5 mg/mL. (d)–(f) Resonator spectra for PEG200, glucose, and BSA solution under a constant concentration of 20 mg/mL. The flow rate ranges from 10 to 90  $\mu\text{L}/\text{min}$  with an increment of 20  $\mu\text{L}/\text{min}$ . (g) For the resonant spectrum of the mixture with 15 mg/mL PEG200, 5 mg/mL glucose, and 15 mg/mL BSA at 50  $\mu\text{L}/\text{min}$  flow rate; (h) for the resonant spectrum of the mixture with 25 mg/mL PEG200, 5 mg/mL glucose, and 5 mg/mL BSA at 50  $\mu\text{L}/\text{min}$  flow rate; and (i) for the resonant spectrum of the mixture with 20 mg/mL PEG200, 10 mg/mL glucose, and 5 mg/mL BSA at 50  $\mu\text{L}/\text{min}$  flow rate.

[Fig. 3(c)] increase from 0 to 30 mg/mL with an increment of 5 mg/mL. In all cases, the shift of the resonant peak increases with time and eventually reaches a plateau. The constant value at the plateau is dependent on the concentration of the chemical—the higher the concentration, the larger the resonant wavelength shifts. The representative resonant spectra of individual components at a constant concentration of 20 mg/mL are shown in Figs. 3(d)–3(f). The spectra of each sample are obtained at five different flow rates ranging from 10 to 90  $\mu\text{L}/\text{min}$  with an increment of 20  $\mu\text{L}/\text{min}$ . As seen in the figures, all measurements for each chemical component reach almost the same plateau regardless of the flow rate, which corroborates that the degree of resonant shift is determined by the chemical concentration. In both the constant-concentration and constant-flow-rate scenarios, samples under different conditions show distinct behaviors in the response region (between zero and the maximum shift). It is speculated that the response region results from the diffusion of target molecules at the interface between the sample solution and running buffer. The microfluidic channel is filled with running buffer before the sample injection. Due to laminar flow at the microscale, the sample solution and the running buffer do not mix at the interface. As the sample solution flows towards the ring resonator, target molecules in the sample solution diffuse into the running buffer and generate a chemical gradient across the interface. The front portion of the chemical gradient, which has a low concentration of target molecules, reaches the ring resonator first, causing a slight shift in resonant wavelength. As the interface flows over the ring resonator, the effective concentration detected by the ring resonator increases with the concentration gradient until it reaches the bulk concentration evidenced by the plateau. The higher the target concentration, the faster the resonant wavelength shifts due to steeper concentration gradient as evidenced by the larger slope in the response regions [Figs. 3(a)–3(c)]. The response region of the resonant spectrum hence contains rich information related to the concentration, diffusion coefficient, and other physical and chemical properties of the target molecules. Similarly, the flow rate also has a strong effect on the response region of the resonant spectrum through the concentration gradient. As the flow rate increases, it takes a shorter time for the interface to reach the ring resonator and hence less time for the molecules to diffuse, causing a steeper concentration gradient at the interface. As a result, the resonant spectrum shows a rapid takeoff and a larger slope in the response region [Figs. 3(d)–3(f)].

The representative resonant spectra of sample mixtures are shown in Figs. 3(g)–3(i). Unfortunately, the resonant spectra of the mixtures do not show prominent features unique to each specific component. Because the key feature regions (i.e., the response regions) of the three components substantially overlap with each other, it is difficult to perform multicomponent chemical analysis by using handcrafted features. Fortunately, we are able to train a neural network for multicomponent chemical analysis using the resonant spectrum of the mixture acquired with the ring resonator. The training of the neural network is described in detail in Section 2.B. The neural network is able to predict the concentration of each individual component based on the resonant spectrum of the mixture.

The accuracy of the neural network model is evaluated with the testing dataset by comparing the predicted composition of the mixture (i.e., the concentrations of all three components) with the ground truth. The prediction error is defined by the RMSE between the predicted values and the ground truth. Each sample composition is measured at five different flow rates, and the predicted compositions at these flow rates are averaged to improve the accuracy of prediction. The neural network model has a low prediction error on the entire testing dataset, ranging only from 0.13 to 2.28 mg/mL over the tested range of 0–30 mg/mL. Figure 4(a) presents an intuitive visualization of the relationship between the predicted composition (magenta dots) and the ground truth (blue dots) of 10 representative mixtures. As seen in the figure, each predicted composition is located close to the respective ground truth, suggesting a high prediction accuracy. The predicted compositions of three mixtures are listed in Table 1. In this case, RMSE is reflected by the distance between the prediction point and ground truth point in Fig. 4(a), which is a measure of difference between the prediction and the ground truth. While the neural network model gives reasonably accurate prediction of the composition at each flow rate, the prediction error may vary over a relatively wide range from one flow rate to another (Tables 1–3). It is noticed that the average prediction values lead to small prediction errors. Hence, the average prediction values are used to determine the composition of the mixture. The distribution of the prediction array of all samples in the testing dataset is shown in Fig. 4(b). All prediction errors are below 2.5 mg/mL over



**Fig. 4.** Composition predicted by the neural network model. (a) 3D distribution of the prediction composition and the ground truth (10 representative points). (b) Histogram of prediction error (RMSE) for all 114 samples in the testing dataset.

**Table 1. Representative Sample 1 Composition Prediction under Different Experiment Conditions<sup>a</sup>**

Flow Rate ( $\mu\text{L}/\text{min}$ )	PEG200 (mg/mL)	Glucose (mg/mL)	BSA (mg/mL)	Prediction Error RMSE (mg/mL)
10	15.4	5.9	14.8	0.58
30	15.7	4.6	15.3	0.50
50	15.0	6.0	14.3	0.70
70	14.6	4.9	14.6	0.33
90	14.3	5.1	14.8	0.42
Average	15.0	5.3	14.8	0.21

<sup>a</sup>Predicted composition of the sample with the expected composition (ground truth) of PEG200 of 15 mg/mL, glucose of 5 mg/mL, and BSA of 15 mg/mL.

**Table 2. Representative Sample 2 Composition Prediction under Different Experiment Conditions<sup>a</sup>**

Flow Rate ( $\mu\text{L}/\text{min}$ )	PEG200 (mg/mL)	Glucose (mg/mL)	BSA (mg/mL)	Prediction Error RMSE (mg/mL)
10	20.7	25.7	15.8	0.73
30	21.5	24.6	15.8	1.01
50	20.4	25.5	14.6	0.44
70	19.9	25.2	13.9	0.65
90	19.5	25.0	14.2	0.54
Average	20.4	25.3	14.9	0.29

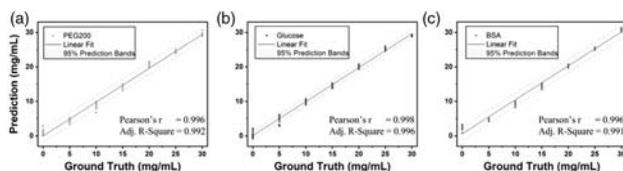
<sup>a</sup>Predicted composition of the sample with the expected composition (ground truth) of PEG200 of 20 mg/mL, glucose of 25 mg/mL, and BSA of 15 mg/mL.

**Table 3. Representative Sample 3 Composition Prediction under Different Experiment Conditions<sup>a</sup>**

Flow Rate ( $\mu\text{L}/\text{min}$ )	PEG200 (mg/mL)	Glucose (mg/mL)	BSA (mg/mL)	Prediction Error RMSE (mg/mL)
10	21.1	10.4	4.1	0.85
30	20.6	10.1	5.3	0.39
50	19.6	11.4	4.8	0.85
70	19.4	10.3	5.4	0.45
90	19.4	9.7	5.3	0.42
Average	20.0	10.4	5.0	0.23

<sup>a</sup>Predicted composition of the sample with the expected composition (ground truth) of PEG200 of 20 mg/mL, glucose of 10 mg/mL, and BSA of 5 mg/mL.

the tested range of 0–30 mg/mL, with more than 75% of the predictions having an error no greater than 1 mg/mL. The median is 0.63 mg/mL, and the average is 0.76 mg/mL with standard deviation of 0.47 mg/mL. The predicted concentrations of the three components (PEG200, glucose, and BSA) for all the samples in the testing dataset are plotted against the ground truth (i.e., the expected concentrations) as shown in Fig. 5. The predicted concentrations show a good linear correlation with the ground truth, with almost all the data points



**Fig. 5.** Predicted concentration versus expected concentration (ground truth) for (a) PEG200, (b) glucose, and (c) BSA.

falling within the 95% prediction bands, again suggesting accurate prediction of our neural network model.

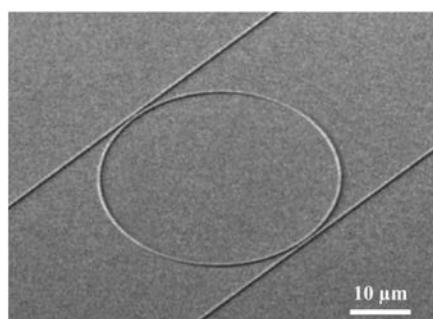
Conventional ring resonator-based sensors only use a single parameter, which is the maximal shift of the resonant wavelength at the steady state (i.e., the plateau in the resonant spectrum), to detect the target. Such a single parameter carries very limited information of the target molecules. By monitoring the resonant wavelength shift over time and plotting the entire resonant spectrum, our sensor is able to acquire more information on the physical and chemical properties of the target molecules. Nevertheless, these pieces of information are buried in the response region of the resonant spectrum because the response regions of individual components substantially overlap with each other. As a result, it is a great challenge to predict the composition of a mixture by decomposing its resonant spectrum using conventional algorithms that are based on handcrafted features. Fortunately, the buried features can be uncovered by machine learning. The machine learning model is able to predict the composition of a mixture with a high accuracy. By combining the highly sensitive ring resonator-based sensor and the intelligent neural network model, we demonstrate a highly accurate smart sensor for label-free multicomponent chemical analysis.

#### 4. CONCLUSIONS

In summary, we demonstrate a machine learning assisted smart sensor for label-free multicomponent chemical analysis. The smart sensor uses the resonant spectrum of a mixture measured by the high-sensitivity and label-free ring resonator to perform multicomponent chemical analysis. The prediction of the mixed compounds is realized by utilizing a neural network model. It shows a high prediction accuracy with a low RMSE ranging only from 0.13 to 2.28 mg/mL. The predicted concentrations of all compounds fall within the 95% prediction bands. With its simple label-free detection strategy and high accuracy, the smart sensor unveils great potential applications for the analysis of various chemical compounds. It also opens up a new avenue for the development of data-driven smart sensors.

#### APPENDIX A: FABRICATION PROCESS OF MICRORING RESONATOR

The sensing structure, a microring resonator, is fabricated on a silicon-on-insulator (SOI) wafer with a 220-nm Si top layer. First, the 70-nm thick silicon is etched by inductively coupled plasma reactive ion etching (RIE) to form grating couplers, and the remaining 150-nm Si is etched to define the wire waveguides. Subsequently, an upper cladding layer of 1.5-μm



**Fig. 6.** SEM of the fabricated microring resonator.

silica is deposited by plasma-enhanced chemical vapor deposition (PECVD). After that, the sensing window of the microring resonator is opened by anisotropic RIE dry etching and isotropic etching based on vapor hydrogen fluoride (HF) to remove the silica upper cladding. The scanning electron micrograph (SEM) of the fabricated microring resonator is shown in Fig. 6.

**Funding.** National Research Foundation Singapore (PUB-1804-0082, NRF-CRP13-2014-01); Ministry of Education—Singapore (MOE2017-T3-1-001).

**Disclosures.** The authors declare no conflicts of interest.

<sup>†</sup>These authors contributed equally to this work.

#### REFERENCES

- E. H. Mordan, J. H. Wade, Z. S. B. Wiersma, E. Pearce, T. O. Pangburn, A. W. deGroot, D. M. Meunier, and R. C. Bailey, "Silicon photonic microring resonator arrays for mass concentration detection of polymers in isocratic separations," *Anal. Chem.* **91**, 1011–1018 (2019).
- R. M. Graybill, C. S. Para, and R. C. Bailey, "PCR-free, multiplexed expression profiling of microRNAs using silicon photonic microring resonators," *Anal. Chem.* **88**, 10347–10351 (2016).
- J. H. Wade, A. T. Alsop, N. R. Vertin, H. Yang, M. D. Johnson, and R. C. Bailey, "Rapid, multiplexed phosphoprotein profiling using silicon photonic sensor arrays," *ACS Cent. Sci.* **1**, 374–382 (2015).
- J. H. Wade and R. C. Bailey, "Applications of optical microcavity resonators in analytical chemistry," *Annu. Rev. Anal. Chem.* **9**, 1–25 (2016).
- Y. Sun and X. Fan, "Optical ring resonators for biochemical and chemical sensing," *Anal. Bioanal. Chem.* **399**, 205–211 (2011).
- C. D. K. Sloan, M. T. Marty, S. G. Sligar, and R. C. Bailey, "Interfacing lipid bilayer nanodiscs and silicon photonic sensor arrays for multiplexed protein-lipid and protein-membrane protein interaction screening," *Anal. Chem.* **85**, 2970–2976 (2013).
- W. W. Shia and R. C. Bailey, "Single domain antibodies for the detection of ricin using silicon photonic microring resonator arrays," *Anal. Chem.* **85**, 805–810 (2013).
- D. Patra and A. Mishra, "Recent developments in multi-component synchronous fluorescence scan analysis," *TrAC Trends Anal. Chem.* **21**, 787–798 (2002).
- A. H. Kamal, S. F. El-Malla, and S. F. Hammad, "A review on UV spectrophotometric methods for simultaneous multicomponent analysis," *Eur. J. Pharm. Med. Res.* **3**, 348–360 (2016).
- S. J. Barton, B. M. Hennelly, T. Ward, K. Domijan, and J. Lowry, "A review of Raman for multicomponent analysis," *Proc. SPIE* **9129**, 91290C (2014).

11. I. Toumi, S. Caldarelli, and B. Torrésani, "A review of blind source separation in NMR spectroscopy," *Prog. Nucl. Magn. Reson. Spectrosc.* **81**, 37–64 (2014).
12. P. Geladi and B. R. Kowalski, "Partial least-squares regression: a tutorial," *Anal. Chim. Acta* **185**, 1–17 (1986).
13. R. Wehrens and B.-H. Mevik, "The pls package: principal component and partial least squares regression in R," *J. Stat. Softw.* **18**, 1–24 (2007).
14. M. C. U. Araújo, T. C. B. Saldanha, R. K. H. Galvão, T. Yoneyama, H. C. Chame, and V. Visani, "The successive projections algorithm for variable selection in spectroscopic multicomponent analysis," *Chemometrics Intellig. Lab. Syst.* **57**, 65–73 (2001).
15. Y. Roggo, P. Chalus, L. Maurer, C. Lema-Martinez, A. Edmond, and N. Jent, "A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies," *J. Pharm. Biomed. Anal.* **44**, 683–700 (2007).
16. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
17. L. Deng and D. J. F. Yu, "Deep learning: methods and applications," *Found. Trends Signal Process.* **7**, 197–387 (2014).
18. J. J. N. Schmidhuber, "Deep learning in neural networks: an overview," *Neural Netw.* **61**, 85–117 (2015).
19. H. M. Robison, P. Escalante, E. Valera, C. L. Erskine, L. Auvil, H. C. Sasieta, C. Bushell, M. Welge, and R. C. Bailey, "Precision immuno-profiling to reveal diagnostic signatures for latent tuberculosis infection and reactivation risk stratification," *Integr. Biol.* **11**, 16–25 (2019).
20. J. Vamathevan, D. Clark, P. Czodrowski, I. Dunham, E. Ferran, G. Lee, B. Li, A. Madabhushi, P. Shah, M. Spitzer, and S. Zhao, "Applications of machine learning in drug discovery and development," *Nat. Rev. Drug Discov.* **18**, 463–477 (2019).
21. F. Cheng and Z. Zhao, "Machine learning-based prediction of drug-drug interactions by integrating drug phenotypic, therapeutic, chemical, and genomic properties," *J. Am. Med. Inform. Assoc.* **21**, e278–e286 (2014).
22. C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert Syst. Appl.* **59**, 235–244 (2016).
23. W. Zhao, A. Bhushan, A. D. Santamaria, M. G. Simon, and C. E. Davis, "Machine learning: a crucial tool for sensor design," *Algorithms* **1**, 130–152 (2008).
24. A. Moraru, M. Pesko, M. Porcius, C. Fortuna, and D. J. Mladenic, "Using machine learning on sensor data," *J. Comput. Inf. Syst.* **18**, 341–347 (2010).
25. M. A. Alsheikh, S. Lin, D. Niyato, and H.-P. Tan, "Machine learning in wireless sensor networks: algorithms, strategies, and applications," *Commun. Surveys Tuts.* **16**, 1996–2018 (2014).
26. Z. Hou, T. Tang, J. Shen, C. Li, and F. Li, "Prediction network of metmaterial with split ring resonator based on deep learning," *Nanoscale Res. Lett.* **15**, 83 (2020).
27. Y. Chen, J. Zhu, Y. Xie, N. Feng, and Q. H. Liu, "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," *Nanoscale* **11**, 9749–9755 (2019).
28. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
29. J. S. T. Smalley, Y. Zhao, A. A. Nawaz, Q. Hao, Y. Ma, I.-C. Khoo, and T. J. Huang, "High contrast modulation of plasmonic signals using nanoscale dual-frequency liquid crystals," *Opt. Express* **19**, 15265–15274 (2011).
30. M. Ian Lapsley, A. Shahrvan, Q. Hao, B. Krishna Juluri, S. Giardinelli, M. Lu, Y. Zhao, I.-K. Chiang, T. Matsoukas, and T. J. Huang, "Shifts in plasmon resonance due to charging of a nanodisk array in argon plasma," *Appl. Phys. Lett.* **100**, 101903 (2012).
31. R. A. Potyrailo, J. E. Brewer, B. Cheng, M. Carpenter, N. M. Houlihan, and A. Kolmakov, "Bio-inspired gas sensing: boosting performance with sensor optimization guided by 'machine learning,'" *Faraday Discuss.* **223**, 161–182 (2020).
32. Z. S. Ballard, D. Shir, A. Bhardwaj, S. Bazargan, S. Sathianathan, and A. Ozcan, "Computational sensing using low-cost and mobile plasmonic readers designed by machine learning," *ACS Nano* **11**, 2266–2274 (2017).
33. X. Feng, G. Zhang, L. K. Chin, A. Q. Liu, and B. Liedberg, "Highly sensitive, label-free detection of 2,4-dichlorophenoxyacetic acid using an optofluidic chip," *ACS Sens.* **2**, 955–960 (2017).
34. Y. Shi, H. Zhao, K. T. Nguyen, Y. Zhang, L. K. Chin, T. Zhu, Y. Yu, H. Cai, P. H. Yap, P. Y. Liu, S. Xiong, J. Zhang, C.-W. Qiu, C. T. Chan, and A. Q. Liu, "Nanophotonic array-induced dynamic behavior for label-free shape-selective bacteria sieving," *ACS Nano* **13**, 12070–12080 (2019).
35. Y. Shi, H. Zhao, L. K. Chin, Y. Zhang, P. H. Yap, W. Ser, C.-W. Qiu, and A. Q. Liu, "Optical potential-well array for high-selectivity, massive trapping and sorting at nanoscale," *Nano Lett.* **20**, 5193–5200 (2020).
36. Z. Li, J. Zou, H. Zhu, B. T. T. Nguyen, Y. Shi, P. Y. Liu, R. C. Bailey, J. Zhou, H. Wang, Z. Yang, Y. Jin, P. H. Yap, H. Cai, Y. Hao, and A. Q. Liu, "Biotoxin photonic sensors with temperature insensitivity using a cascade of ring resonator and Mach-Zehnder interferometer," *ACS Sens.* **5**, 2448–2456 (2020).
37. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "TensorFlow: a system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation* (2016), pp. 265–283.
38. H. Zhang, M. F. Karim, S. Zheng, H. Cai, Y. Gu, S. S. Chen, H. Yu, and A. Q. Liu, "A high-resolution dual-microring-based silicon photonic sensor using electronic integrated circuit," in *CLEO: Applications and Technology* (Optical Society of America, 2018), paper ATh4O.4.
39. H. Zhang, M. F. Karim, S. Zheng, H. Cai, Y. Gu, S. S. Chen, H. Yu, and A. Q. Liu, "Machine learning and silicon photonic sensor for complex chemical components determination," in *CLEO: Science and Innovations* (Optical Society of America, 2018), paper JW2A.54.
40. M. C. Cardenosa-Rubio, H. M. Robison, and R. C. Bailey, "Recent advances in environmental and clinical analysis using microring resonator-based sensors," *Curr. Opin. Environ. Sci. Health* **10**, 38–46 (2019).

# PHOTONICS Research

## Monte Carlo simulation fused with target distribution modeling via deep reinforcement learning for automatic high-efficiency photon distribution estimation

JIANHUI MA,<sup>1</sup> ZUN PIAO,<sup>1</sup> SHUANG HUANG,<sup>1</sup> XIAOMAN DUAN,<sup>1</sup> GENGGENG QIN,<sup>1</sup> LINGHONG ZHOU,<sup>1,2</sup> AND YUAN XU<sup>1,3</sup>

<sup>1</sup>School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China

<sup>2</sup>e-mail: smart@smu.edu.cn

<sup>3</sup>e-mail: yuanxu@smu.edu.cn

Received 26 October 2020; revised 17 December 2020; accepted 21 December 2020; posted 21 December 2020 (Doc. ID 413486); published 8 February 2021

Particle distribution estimation is an important issue in medical diagnosis. In particular, photon scattering in some medical devices extremely degrades image quality and causes measurement inaccuracy. The Monte Carlo (MC) algorithm is regarded as the most accurate particle estimation approach but is still time-consuming, even with graphic processing unit (GPU) acceleration. The goal of this work is to develop an automatic scatter estimation framework for high-efficiency photon distribution estimation. Specifically, a GPU-based MC simulation initially yields a raw scatter signal with a low photon number to hasten scatter generation. In the proposed method, assume that the scatter signal follows Poisson distribution, where an optimization objective function fused with sparse feature penalty is modeled. Then, an over-relaxation algorithm is deduced mathematically to solve this objective function. For optimizing the parameters in the over-relaxation algorithm, the deep  $Q$ -network in the deep reinforcement learning scheme is built to intelligently interact with the over-relaxation algorithm to accurately and rapidly estimate a scatter signal with the large range of photon numbers. Experimental results demonstrated that our proposed framework can achieve superior performance with structural similarity  $>0.94$ , peak signal-to-noise ratio  $>26.55$  dB, and relative absolute error  $<5.62\%$ , and the lowest computation time for one scatter image generation can be within 2 s. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.413486>

### 1. INTRODUCTION

Particle distribution estimation is an important issue in medical diagnosis because the principle of that distribution in some crucial imaging devices follows the law of large numbers [1]. However, particle distribution estimation by only trillions of particle sampling is extremely time-consuming. Thus, approximating a target function model on the basis of target distribution feature analysis can substantially hasten particle estimation [2]. Particles fused with a target distribution feature for a few particles have a promising prospect in rapid statistical estimation without the loss of accuracy. As an important branch of particle distribution estimation, photon scattering is commonly involved in X-ray imaging, especially cone beam computed tomography (CBCT), which has been widely used in medical imaging given its high spatial resolution and low radiation dose and for applications such as dental CBCT, image-guided radiotherapy, and extremity CBCT [3]. Scatter occurs when photons

emitted from an X-ray source interact with a physical object and can extremely degrade the reconstructed image quality for clinical diagnosis, thereby leading to pixel-value inaccuracy. Various methods have been proposed for scatter removal, including the scatter kernel [4,5], beam stop array technique [6,7], and primary modulator [8]. Nevertheless, these approaches require additional equipment or increase the irradiation dose to the patient. The Monte Carlo (MC) particle sampling approach has been proved to estimate accurate scatter signals in a cost-efficient manner without the above disadvantages [9–11]. The said technique can precisely simulate the physical process of photon transport to produce all types of scatters composed of Compton, Rayleigh, and photoelectric effect. In reality, one X-ray tube generates approximately  $1 \times 10^{13}$  source photons for imaging one projection under one angle view. Therefore, the scatter simulation of photon transport by only MC particle sampling is extremely time-consuming, and this feature precludes its clinical application [12].

To enhance its computational efficiency, the scatter signal is typically simulated by an MC algorithm with a low photon number, so the generated signal is noisy and fragmentary. To overcome this problem, the noise-contaminated scatter signal is assumed to be governed by the Poisson distribution and sparse feature representation. An efficient denoising algorithm based on Poisson distribution and sparse feature fusion is proposed to smooth the signal along the spatial dimension. Nonetheless, the weight parameter between the Poisson distribution and sparse feature does not follow a one-size-fits-all trend for all cases with different photon numbers. Tuning parameters manually for all situations is a trial-and-error process, which calls for considerable time and effort. Moreover, that scheme hinders a noise removal algorithm from reaching an optimal solution of generating an accurate scatter signal via ultra-low photons.

In the past few years, deep-learning algorithms have achieved remarkable success across many fields, such as computer vision and pattern recognition [13–18]. Deep-learning methods build deep neural networks (DNN) to directly extract local and global features from the dataset, an approach that can avoid handcrafted selection [13]. Motivated by the powerful performance of deep learning, the Google DeepMind group integrated a neural network and reinforcement learning (RL) algorithm to play Atari games with human-level performance [19]. Afterwards, AlphaGo based on deep RL (DRL) defeated human masters in the ancient Chinese game Go and attracted global attention [20]. Shen *et al.* [21] proposed a parameter tuning policy network to adjust pixel-wise parameters in iterative computed tomography (CT) reconstruction and attain comparable or better image quality. DRL was also applied in radiotherapy for optimal dose adaptation and a treatment planning optimization problem [22–25]. These superior performances have demonstrated that the DRL algorithm can achieve a task analogous to human instincts. In DRL, an agent represented by a DNN interacts with the environment to explore all possible consequences of the action for the highest reward feedback. The agent acts according to their observation of the environment, which, in turn, is changed by the action and yields a new observation for the next step. Over many episodes, the agent is supposed to develop an optimal policy to obtain maximum rewards. Essentially, parameter tuning is a dynamic decision-making process, for which DRL is highly suitable. Inspired by the impressive achievements of pioneering work and the rationale of DRL, we propose a framework to realize automatic high-efficiency scatter estimation via the fusion of MC particle sampling, statistical distribution features, and a DRL scheme.

The remainder of this paper is organized as follows. Section 2 describes the automatic scatter estimation framework (ASEF), including its key components, namely, the MC particle sampling algorithm, Poisson distribution and sparse feature-based statistical distribution algorithm, and a DRL scheme. Network interpretability and implementation details are also introduced in the section. Section 3 evaluates the performance of the proposed framework qualitatively and quantitatively. Section 4 discusses and summarizes the strengths and drawbacks.

## 2. METHODS AND MATERIALS

### A. Automatic Scatter Estimation Framework

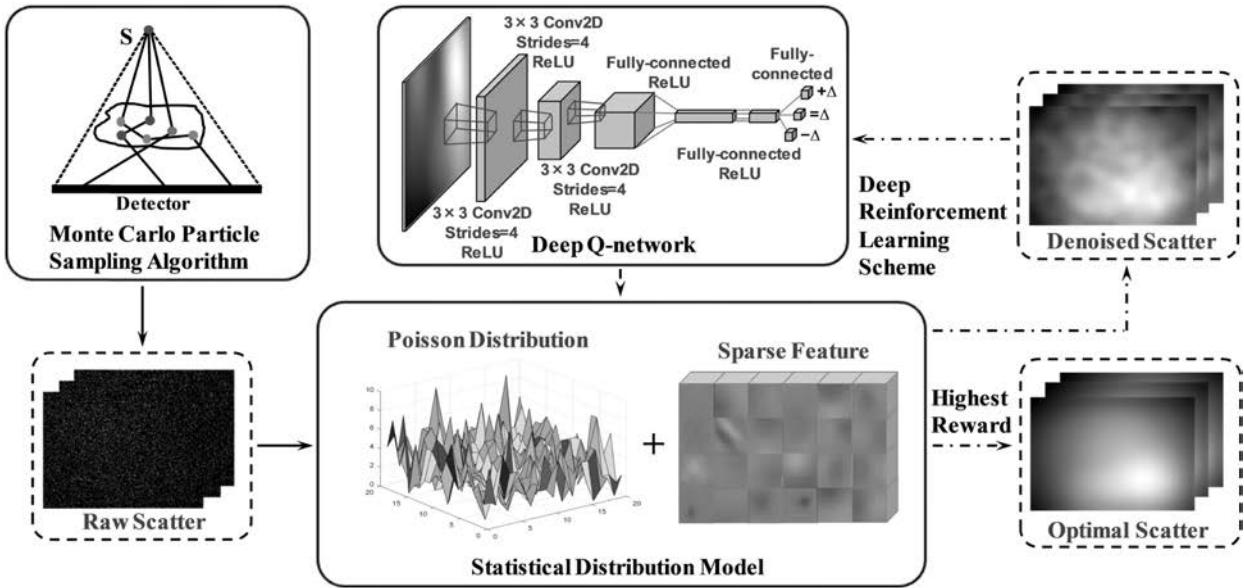
In this study, we propose an ASEF that integrates an MC particle sampling algorithm (details depicted in Section 2.B), a statistical distribution model fused with sparse feature representation under the Poisson distribution assumption (Section 2.C), and a DRL scheme (Section 2.D). The entire framework is illustrated in Fig. 1. First, on the basis of the X-ray source energy spectrum and system geometry configuration, a graphic processing unit (GPU)-based MC simulation initially yields a raw scatter signal with a low photon number to hasten scatter generation. Second, assuming that the scatter signal follows Poisson distribution, an optimization objective function fused with sparse feature penalty is constructed. Then, an over-relaxation algorithm is deduced mathematically to solve this objective function. For optimizing the parameters in the over-relaxation algorithm, a neural network called the deep *Q*-network (DQN) is built to intelligently interact with the over-relaxation algorithm for optimal scatter image quality. Specifically, the well trained DQN, which possesses the optimal policy of the highest rewards, takes an action to change the parameters; then the over-relaxation algorithm yields a scatter image based on the adjusted parameters for the next action selection. After several tuning steps, the optimal parameters can be determined when the highest cumulative rewards of a sequence of steps are obtained. The scatter image yielded by the over-relaxation algorithm with the optimal parameters is referred to as optimal quality.

### B. Monte Carlo Particle Sampling Algorithm

An in-house MC simulation tool with a polychromatic energy spectrum is utilized in this study for the MC particle sampling of the scatter signal. The energy spectrum describes the probability density of a source photon as a function of its energy. Such an energy spectrum can be specified by using the method developed by Boone *et al.* [26]. According to tabulated data including the attenuation coefficients and the cross sections of Compton scattering, Rayleigh scattering, and the photoelectric effect, each photon transporting through the imaging phantom is simulated to compute the particle distribution in the virtual detector. During the simulation, an indicator is carried by each photon that records if any scattering events have taken place during the transport. For different scattering events of different orders, the photon will be marked by a number as the indicator of the photon. For example, the indicators of all primal photons are zero, and all first order Compton photons are marked one as the indicator. The CT volume of the phantom is utilized to generate different material and density information on the basis of the CT value via threshold-based segmentation. The MC sampling scheme includes several realistic features, e.g., the polychromatic source spectrum, materials and thicknesses of the detector, and beam collimation.

In the MC particle sampling, the scatter angles  $\theta$  for Rayleigh scattering are sampled from the expression for the differential cross sections as

$$\frac{d\sigma_R}{d(\cos \theta)} = \pi r_0^2 F(E, \theta, Z)^2 (1 + \cos^2 \theta), \quad (1)$$



**Fig. 1.** Automatic scatter estimation framework. The MC algorithm generates raw scatter signals in terms of the X-ray source energy spectrum and system geometry configuration. The DRL scheme (denoted by the dashed black arrow) employs a deep *Q*-network to interact with the statistical distribution model to yield a satisfactory scatter image.

and the scatter angles  $\theta$  for Compton scattering can be expressed as

$$\frac{d\sigma_C}{d(\cos \theta)} = \pi r_0^2 S(E, \theta, Z) [P(E, \theta) - P(E, \theta)^2 \sin^2 \theta + P(E, \theta)^3]. \quad (2)$$

Note that the photon energy is reduced after the Compton scattering, and  $P(E, \theta)$  is the ratio of the energy  $E'$  before scattering and the current energy  $E$ :

$$P(E, \theta) = \frac{E'}{E} = \frac{1}{1 + \frac{E}{m_0 c^2} (1 - \cos \theta)}, \quad (3)$$

where  $r_0$ ,  $c$ , and  $m_0$  separately indicate the electron radius, light speed, and resting mass of one electron in Eqs. (1)–(3). The functions  $F(E, \theta, Z)$  and  $S(E, \theta, Z)$  are the form factor and the incoherent scattering function. Data for  $F(E, \theta, Z)$  and  $S(E, \theta, Z)$  are generated from the gCTD package [27], which is a fast MC simulation package for patient-specific CBCT imaging or dose calculation. For shortening computation time, scatter angles  $\theta$  were not sampled by calculating Eqs. (1) and (2) but were taken from the pre-calculated look-up table. Electron transport is neglected in the simulation, as it cannot reach the detector and will be absorbed in the phantom immediately. Moreover, the Woodcock tracking algorithm [28] is employed to avoid calculating the integral of photon path attenuation voxel by voxel.

### C. Scatter Statistical Distribution Model Based on Poisson Distribution and Sparse Feature Representation

It has been proved that the measurement determined by the photon distribution statistics follows a Poisson distribution [29,30]. The scatter signal  $\hat{S}(u)$ , which is noise-contaminated, is expected to approach true scatter  $S(u)$ , and thus its probability density function with expectation  $S(u)$  is defined as

$$P[x = \hat{S}(u)] = \frac{S(u)^{\hat{S}(u)}}{\hat{S}(u)!} e^{-S(u)}, \quad (4)$$

where  $u = (i, j)$  indicates the detector coordinate, for which  $i$  and  $j$  denote the horizontal and vertical axes, respectively. Note that measurements  $\hat{S}(u)$  are independent of each other, and thus

$$P(x = \hat{S}) = \int \frac{S^{\hat{S}}}{\hat{S}!} e^{-S} d\mu. \quad (5)$$

Moreover, Eq. (5) is maximal when  $\hat{S} \approx S$ , and we minimize  $-\log P(x = \hat{S})$  rather than maximizing  $P(x = \hat{S})$ . That is

$$\arg \min_S \int (S - \hat{S} \log S) d\mu. \quad (6)$$

Given the ill-posed nature of Eq. (6), a total-variation (TV) regularization [31] is employed for the smoothness of the solution. Therefore, the formula based on Poisson distribution and sparse feature representation can be defined as follows:

$$S(u) = \arg \min_S \int (S - \hat{S} \log S) d\mu + \frac{\beta}{2} \int |\nabla S|^2 d\mu, \quad (7)$$

where the first data-fidelity term is designed to ensure Poisson distribution, and the second TV regularization term is formulated for noise suppression.  $\beta$  is a hyper-parameter controlling the relationship between the data-fidelity and regularization terms. As Eq. (7) is convex, the derivative of its optimal solution is equal to zero, that is,

$$\left(1 - \frac{\hat{S}}{S}\right) - \beta \nabla^2 S = 0. \quad (8)$$

Equation (8) can be discretized as

$$\left[1 - \frac{\hat{S}(i,j)}{S(i,j)}\right] - \beta \nabla^2 S(i,j) = 0, \quad (9)$$

where the Laplacian operator  $\nabla^2 S(i,j) = \sum S(i,j) - 4S(i,j)$ , and  $\sum S(i,j)$  stands for  $S(i-1,j) + S(i+1,j) + S(i,j-1) + S(i,j+1)$ . Then, an over-relaxation iteration algorithm is employed to solve Eq. (7):

$$\begin{aligned} S^{(k+1)}(i,j) &= (1-\omega)S^{(k)}(i,j) \\ &+ \frac{\omega}{4} \left[ \sum S^{(k)}(i,j) - \frac{1}{\beta} \left( 1 - \frac{\hat{S}(i,j)}{S^{(k)}(i,j)} \right) \right], \end{aligned} \quad (10)$$

where the superscript  $k$  represents the iteration step, and  $\omega$  is an empirical constant to speed up convergence. Finally, the noise free scatter signal  $S$  can be obtained by solving Eq. (10) iteratively.

## D. Deep Reinforcement Learning

### 1. Double Deep Q-Network

Three hyper-parameters appear in Eq. (10):  $\omega$ ,  $\beta$ , and the number of iterations  $k$  to be determined. Typically, these variables are set to a constant to obtain smoothed scatter images even though the simulated photon number is different. Tuning parameters manually for all situations is a trial-and-error process, which calls for considerable time and effort. Moreover, this approach restricts the over-relaxation algorithm from exploring a global optimal solution for improved image quality. These quality-related parameters are not one-size-fits-all for all cases with different photon numbers. Consequently, we aim to utilize a DRL scheme to intelligently seek out satisfactory images like human behavior.

The  $Q$ -learning algorithm is a widely used approach, for which the action-value function  $Q_\pi(s, a)$  is defined as the expectation of the rewards sum of all possible steps from the current state  $s$  taking action  $a$  under policy  $\pi$ :

$$Q_\pi(s, a) = E(G_t | s_t = s, a_t = a), \quad (11)$$

$$G_t = r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \dots = \sum_{m=1}^{\infty} \gamma^m r_{t+m}, \quad (12)$$

where  $G_t$  is the discounted rewards sum starting from step  $t$ ,  $r_t$  is the reward at step  $t$ , and  $\gamma \in [0, 1]$  is a discount factor. The  $Q$ -learning algorithm aims to explore an optimal policy to maximize the action-value function:

$$Q_\pi^*(s, a) = \max_{\pi} Q_\pi(s, a). \quad (13)$$

The optimal action-value function  $Q_\pi^*(s, a)$  can be solved iteratively via the Bellman equation:

$$Q_\pi^*(s, a) = r + \gamma \max_{a'} Q_\pi^*(s', a'), \quad (14)$$

where  $r$  denotes the reward after adopting action  $a$  to the current state  $s$ , and  $s'$  is the next state following the current state  $s$  under action  $a$ . Equation (14) entails expensive computation when the state and action involve extensive dimensions. Thus, we adopt a convolutional neural network (CNN)  $Q_\pi(s, a; W)$  to approximate the  $Q$  function  $Q_\pi(s, a)$ , such that a quadratic loss function defined as Eq. (15) can be minimized to force  $Q_\pi(s, a; W)$  to approach optimal  $Q_\pi^*(s, a; W)$ :

$$L = E\{[r + \gamma \max_{a'} Q_\pi(s', a'; W) - Q_\pi(s, a; W)]^2\}. \quad (15)$$

To enhance the training stability of  $Q_\pi(s, a; W)$ , a separate target network  $\hat{Q}_\pi(s, a; \hat{W})$ , whose architecture is the same as the main network  $Q_\pi(s, a; W)$ , is constructed following the protocol in Ref. [19]. The weights  $\hat{W}$  of  $\hat{Q}_\pi(s, a; \hat{W})$  will be fixed when the main network weight  $W$  is updated according to the gradient of loss function defined as

$$L = E\{[r + \gamma \max_{a'} \hat{Q}_\pi(s', a'; \hat{W}) - Q_\pi(s, a; W)]^2\}. \quad (16)$$

After several training steps,  $\hat{W}$  is updated with  $W$ ; that is,  $\hat{W} = W$ . The approach based on the  $Q$ -learning algorithm and DNN is referred to as DQN. The DQN tends to overestimate the action-value function, so we use double DQN (DDQN) for improved robustness [32]. The DDQN utilizes the main network  $Q_\pi(s, a; W)$  to select the action corresponding to the maximum  $Q$ -value, which is the output value of the network, namely,  $a' = \arg \max_{a'} [Q_\pi(s', a'; W)]$ . Then, action  $a'$  will be utilized in the target network  $\hat{Q}_\pi(s', a'; \hat{W})$  to predict the  $Q$ -value. That is,

$$L = E\{[r + \gamma \hat{Q}_\pi(s', \arg \max_{a'} [Q_\pi(s', a'; W)]; \hat{W}) - Q_\pi(s, a; W)]^2\}. \quad (17)$$

### 2. Reward Function

In DRL, the agent aims to take step-by-step action toward the desired situation by obtaining highest rewards. Therefore, the reward is supposed to faithfully evaluate the quality of state  $s$ . In this study, state  $s$  represents the scatter image per angle. Consequently, we employ structural similarity (SSIM) [33] to construct the reward function. The SSIM is a perceptual metric that can quantitatively measure the similarity between two images  $x$  and  $y$ , focuses on luminance, contrast, and structure, and is defined as

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad (18)$$

where  $\mu_x$  and  $\mu_y$  separately denote the average of  $x$  and  $y$ ,  $\sigma_x^2$  and  $\sigma_y^2$  represent variance, and  $\sigma_{xy}$  is the covariance of  $x$  and  $y$ . A higher SSIM indicates greater similarity between two images. Therefore, the reward function consists of the SSIM of the state and ground truth formulated as

$$r_t = \text{sgn}[\text{SSIM}(s_{t+1}, s_{gt}) - \text{SSIM}(s_t, s_{gt})], \quad (19)$$

where  $s_{gt}$  denotes the ground truth of the scatter, and the first and second terms separately measure the SSIM of the scatter at step  $t$  with the ground truth and the SSIM between the scatter at step  $t+1$  and ground truth.  $r_t$  will be positive if  $s_{t+1}$  is closer to the ground truth compared with the previous state  $s_t$ , such that it can inspire the DDQN to improve the generated scatter quality toward the desired image quality.  $\text{sgn}$  indicates the sign function, which transfers a positive number to 1 and a negative number to -1. As suggested in Ref. [19], the sign function is applied to rescale the reward for the scale limitation of the error derivatives.

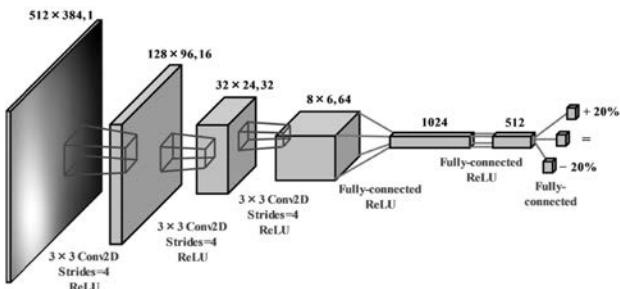
### 3. Training Process of the DDQN

In this study, action  $a$  has three possible actions: increase or decrease the parameter by 20% and keep it invariant. Some

researches [21,23] specified the amount of parameter change with different amplitudes. Because the reward function defined in Eq. (19) utilizes the sign function to restrict the scale of the error derivatives, the network cannot differentiate between rewards of different magnitudes. Therefore, we arbitrarily select 20% as the parameter amplitude. We build our network  $W = \{W_k, W_\omega, W_\beta\}$  containing three subnetworks to represent three parameters in the over-relaxation algorithm defined as Eq. (10). These subnetworks  $\{W_k, W_\omega, W_\beta\}$  possess the same architecture illustrated in Fig. 2. We repeatedly train  $\{W_k, W_\omega, W_\beta\}$  for  $N_{\text{episode}}$  times, and each episode has a series of steps. At step  $t$ , one of the three subnetworks  $\{W_k, W_\omega, W_\beta\}$  is randomly selected with equal probability. Then, the  $\epsilon$ -greedy algorithm is adopted to select an action to adjust the parameter. Specifically, we select a random action  $a_t$  with probability  $\epsilon$ ; otherwise, the action corresponding to the highest  $Q$ -value is selected, i.e.,  $a_t = \arg \max_a [Q_\pi(s_t, a; W)]$ . Afterward, the parameter is adjusted according to  $a_t$ , and the over-relaxation algorithm is solved to generate a new scatter  $s_{t+1}$  as the state for the next step. Equation (19) takes  $s_t$  and  $s_{t+1}$  to calculate the reward  $r_t$ . The dataset  $\{s_t, a_t, r_t, s_{t+1}\}$  is then stored in the experience replay memory  $D$  to mitigate the correlation between the training dataset generated in successive steps. Subsequently, network  $W$  is trained by randomly sampling a mini-batch of dataset from  $D$ .  $\{W_k, W_\omega, W_\beta\}$  will be updated via minimizing loss function in Eq. (17). Finally, update target network weights  $\{\hat{W}_k, \hat{W}_\omega, \hat{W}_\beta\}$  with  $\{W_k, W_\omega, W_\beta\}$ , i.e., let  $\{\hat{W}_k, \hat{W}_\omega, \hat{W}_\beta\} = \{W_k, W_\omega, W_\beta\}$  every  $N_{\text{update}}$  steps. The training process of the DDQN is summarized in Table 1, and the related parameter values are defined in Table 2.

### E. DDQN Interpretability

The DDQN can predict an action under the optimal policy. However, the mechanism by which the DDQN takes action in terms of the current state remains unclear. To interpret the DDQN, a gradient-weighted class activation map (Grad-CAM) [34] is utilized to highlight crucial regions in the state for action prediction. The Grad-CAM can yield a localization map using the gradient of output with respect to the last



**Fig. 2.** Network architecture in the DDQN. The network takes a scatter image as input and predicts three possible actions for parameter adjustment. The number at the top denotes the feature map size and channel number, and the operations for each layer are presented at the bottom. For instance, the first hidden layer convolves 16 filters of  $3 \times 3$  with stride four with the input layer followed by a rectified linear unit (ReLU) activation function, and the output layer is a fully connected linear layer with three outputs.

**Table 1. DDQN Training Process**

1. Initialize main network weights  $W$  and target network weights  $\hat{W}$
2. **For** episode = 1, 2, ...,  $N_{\text{episode}}$  **do**
3.   **For** projection = 1, 2, ...,  $N_{\text{prj}}$  **do**
4.     Initialize  $\{k_0, \omega_0, \beta_0\}$
5.     Generate  $s_1$  using Eq. (10) with  $\{k_0, \omega_0, \beta_0\}$
6.     **For**  $t = 1, 2, \dots, N_{\text{step}}$  **do**
7.       Randomly select one subnetwork from  $\{W_k, W_\omega, W_\beta\}$
8.       With probability  $\epsilon$  select action  $a_t$  randomly
9.       Otherwise choose  $a_t = \arg \max_a [Q_\pi(s_t, a; W)]$
10.      Adjust parameters  $\{k_t, \omega_t, \beta_t\}$  according to  $a_t$
11.      Generate  $s_{t+1}$  using Eq. (10) with  $\{k_t, \omega_t, \beta_t\}$
12.      Compute reward  $r_t$  using Eq. (19)
13.      Store dataset  $\{s_t, a_t, r_t, s_{t+1}\}$  in experience replay  $D$
14.      Randomly sample a mini-batch of dataset from  $D$
15.      Compute the gradient of loss function in Eq. (17)
16.      Update main network weights  $W = \{W_k, W_\omega, W_\beta\}$
17.      For every  $N_{\text{update}}$  steps, let  $\hat{W} = W$
18.     **End For**
19.   **End For**
20. **End For**

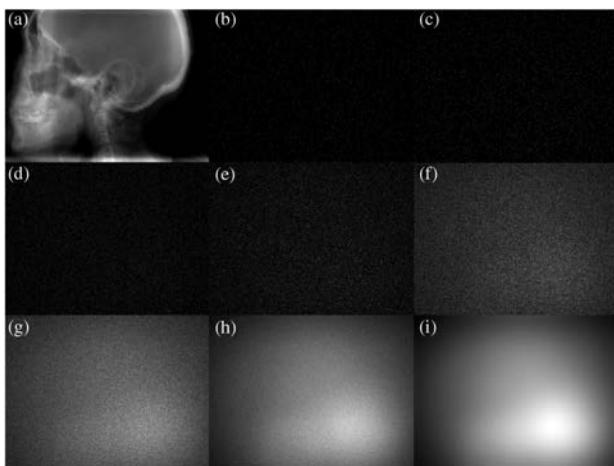
**Table 2. Parameters in the DDQN Training Phase**

Parameters	Values	Descriptions
$N_{\text{episode}}$	100	Number of training episodes
$N_{\text{prj}}$	45	Number of training projections
$N_{\text{step}}$	30	Number of steps for each episode
$N_{\text{update}}$	20	Number of steps for target network weights update
$D$	2000	Capacity of experience replay memory
$\epsilon$	[0.01, 1]	Probability of random action in $\epsilon$ -greedy algorithm
$\gamma$	0.6	Discount factor
$l_r$	0.001	Learning rate of gradient descent for main network
$N_{\text{batch}}$	64	Mini-batch samples for network training

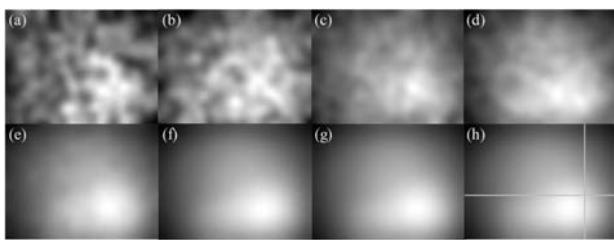
convolution layer to locate regions that are much more important for DDQN decision making. More concretely, Grad-CAM initially calculates the gradient of the  $Q$ -value  $q^a$  with respect to the feature map  $A^k$  of the last convolutional layer whose width and height are separately indexed by  $i$  and  $j$ , namely,  $\frac{\partial q^a}{\partial A_{ij}^k}$ . Then, these gradients are backpropagated and global-average-pooled to attain weights  $\alpha_k^a$  for every feature map, i.e.,  $\alpha_k^a = \frac{1}{Z} \sum_i \sum_j \frac{\partial q^a}{\partial A_{ij}^k}$ . The Grad-CAM heatmap is the sum of the weighted feature maps and followed by a rectified linear unit (ReLU):  $L_{\text{Grad-CAM}}^a = \text{ReLU}(\sum_k \alpha_k^a A^k)$ . Finally, the Grad-CAM heatmap is upsampled to the input state size.

### F. Implementation Details

In this study, we utilize one Nvidia TITAN Xp GPU and four Intel Core i7 3.6 GHz processors with 40 GB memory to implement the framework using TensorFlow [35]. Equation (10) is a noise removal algorithm in the projection domain. Therefore, the scatter  $S$  is a projection whose resolution is  $512 \times 384$ . Six groups of scatter datasets are present with



**Fig. 3.** (a) is the primary projection of the head and neck (H&N) patient; (b)–(i) represent raw scatter projections that are separately calculated by the MC particle sampling algorithm with source photons of  $5 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$ ,  $1 \times 10^9$ ,  $1 \times 10^{10}$ , and  $1 \times 10^{12}$  for the same projection angle.



**Fig. 4.** (a)–(g) are the scatter images of Figs. 3(b)–3(h) smoothed by the over-relaxation smoothing algorithm; (h) corresponds to Fig. 3(i), which is considered a noise free scatter image and utilized as the ground truth.

different photon numbers of  $5 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$ , and  $1 \times 10^9$ . Each dataset contains 90 projections. As shown in Fig. 3(i), a raw scatter signal generated from MC simulation with  $1 \times 10^{12}$  particles is almost noise free and taken as the ground truth. We randomly select 45 of 90 scatter images at  $1 \times 10^6$  for training. The 45 images whose angles

differ from those of the training cases in the six groups are chosen as the testing cases. Therefore, the total number of testing cases is 270.

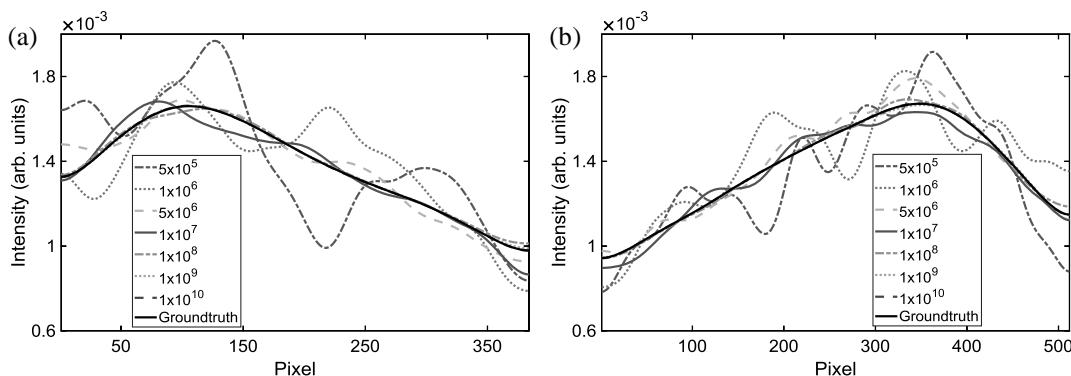
### 3. RESULTS

#### A. MC Particle Sampling Algorithm

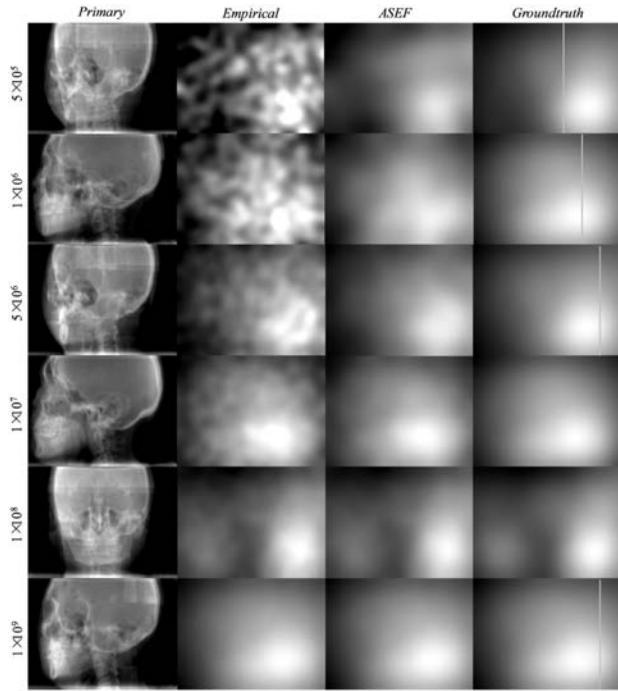
Figures 3(b)–3(i) present the raw scatter signals of the projection produced by the MC particle sampling algorithm with different photon numbers. Figure 3(a) is the primary signal of the projection generated by a typical 100 kVp poly-energetic spectrum with about 100 energy channels. Clearly, Figs. 3(b)–3(h) severely suffer from noise contamination because of the deficiency of source photons (less than  $1 \times 10^{12}$ ), whereas Fig. 3(i) shows the MC algorithm that can precisely yield a desirable quality scatter image after the  $1 \times 10^{12}$  photons was simulated. The quantitative evaluation of image similarity and time cost is shown in Tables 3 and 4, respectively. As presented in Table 4, simulating one projection via MC with  $1 \times 10^{11}$  photons costs 6402.60 s. The computation time of the MC particle sampling algorithm exponentially increases with the growth of the photon number. Therefore, directly producing an approved scatter signal through the MC algorithm would be impractical.

#### B. Over-Relaxation Smoothing Algorithm for Scatter Estimation

Figures 4(a)–4(g) are the corresponding smoothed scatter signals of Figs. 3(b)–3(h) manipulated by an over-relaxation smoothing algorithm with empirical parameters ( $k = 700$ ,  $\omega = 0.8$ , and  $\beta = 1 \times 10^{-11}$ ), which is denoted as *Empirical* for simplicity. The raw scatter projection with  $1 \times 10^{12}$  photons, as shown in Fig. 4(h), is employed as the ground truth for comparison. The over-relaxation smoothing algorithm can improve scatter image quality among all cases compared to the corresponding raw scatter images. Figures 4(f) and 4(g) are visually similar to Fig. 4(h). Figure 4(e) is slightly rough, whereas Figs. 4(a)–4(d) are severely noisy and fragmentary. The intensity profiles of Figs. 4(a)–4(h) along the row and column indicated by an orange line in Fig. 4(h) are plotted to further validate the performance of the over-relaxation smoothing algorithm. As shown in Figs. 5(a) and 5(b), the intensity profiles of Figs. 4(f) and 4(g) follow exactly the same trend as that in Fig. 4(h). Moreover, the profile lines of Figs. 4(a)–4(e) still



**Fig. 5.** Intensity profiles of Fig. 4 along the (a) horizontal and (b) vertical directions as denoted by the orange lines in Fig. 4(h).



**Fig. 6.** From top to bottom: six testing results with  $5 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ ,  $1 \times 10^7$ ,  $1 \times 10^8$ , and  $1 \times 10^9$  source photons. From left to right: primary signals, smoothed scatter signals restored by the over-relaxation algorithm with empirical parameters, smoothed scatter signals restored by the proposed framework, and the ground truth.

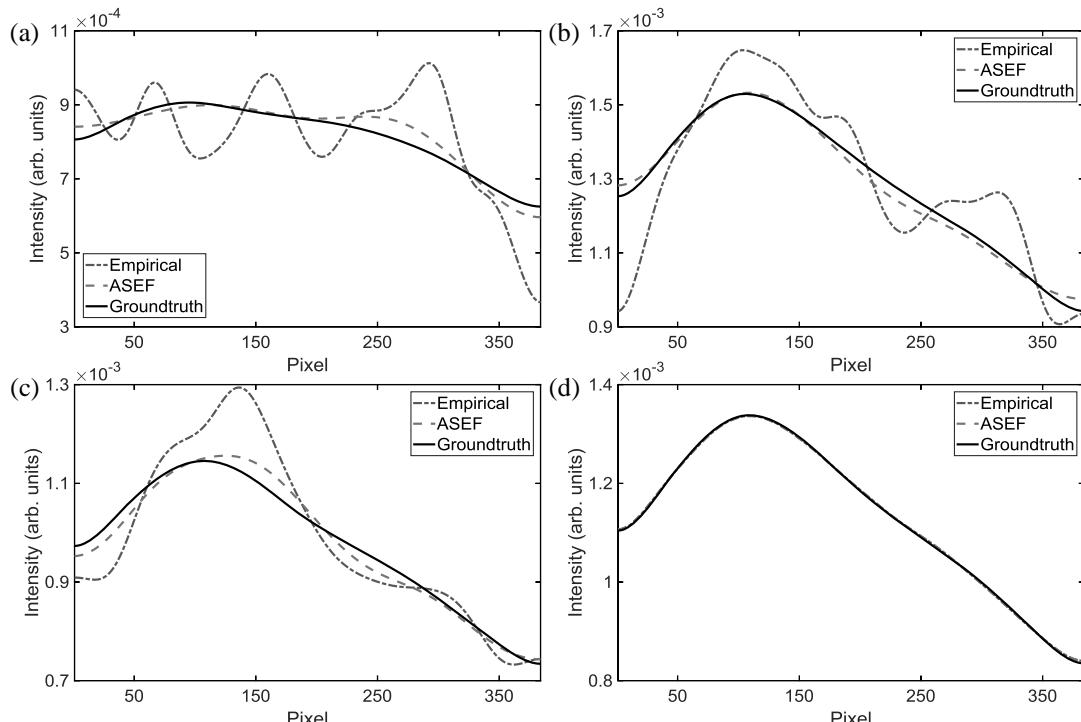
oscillate. Therefore, the over-relaxation smoothing algorithm performs well when the photon number is no less than  $1 \times 10^9$ .

### C. Scatter Estimation Performance Comparison

To validate its scatter estimation performance, the proposed framework named *ASEF* is compared with *Empirical* among six testing groups with dissimilar photon numbers. As depicted in Fig. 6, *Empirical* alleviates the noise component to some extent but remains inferior compared to *ASEF*. Although both *Empirical* and *ASEF* results degrade as the number of source photons decreases, *ASEF* scatter images are much more consistent in quality improvement. For the most challenging case with  $5 \times 10^5$  source photons in the first row of Fig. 6, the *Empirical* scatter image is severely distorted after the over-relaxation smoothing algorithm. By contrast, the *ASEF* can substantially remove noise and restore the true scatter signal from raw scatter. The *ASEF* scatter image is visually smoother and similar to the ground truth.

The intensity profiles of Fig. 6 along the vertical direction exhibit the scatter variation trend shown in Fig. 7. Figures 7(a)–7(c) indicate that the *ASEF* can follow a similar trend to the ground truth, whereas the profile of *Empirical* oscillates diversely because of the existence of noise. Figure 7(d) suggests that both *Empirical* and *ASEF* can obtain satisfactory scatter image quality for the  $1 \times 10^9$  case. However, *ASEF* possesses a powerful capability of exploring optimal solutions for different source photons (Figs. 6 and 7).

For quantitative evaluation, the SSIM, peak signal-to-noise ratio (PSNR), and relative absolute error (RAE) are employed, in which PSNR and RAE are defined as



**Fig. 7.** (a)–(d) Intensity profiles of the first, second, third, and last rows in Fig. 6. The locations of the profiles (a)–(d) are denoted by orange lines at the last column of Fig. 6.

**Table 3.** SSIM, PSNR, and RAE Statistics (avg.  $\pm$  std.) among All Testing Cases<sup>a</sup>

Photon Number	SSIM (1 = Best)				PSNR (dB)				RAE (%)			
	<i>Empirical</i>		<i>ASEF</i>		<i>Empirical</i>		<i>ASEF</i>		<i>Empirical</i>		<i>ASEF</i>	
	avg.	std.	avg.	std.	avg.	std.	avg.	std.	avg.	std.	avg.	std.
$5 \times 10^5$	0.79	$4.70 \times 10^{-2}$	<b>0.94</b>	$2.36 \times 10^{-2}$	21.54	0.85	<b>26.55</b>	<b>1.34</b>	12.03	$1.27 \times 10^{-2}$	<b>5.62</b>	$1.27 \times 10^{-2}$
$1 \times 10^6$	0.88	$3.73 \times 10^{-2}$	<b>0.96</b>	$1.67 \times 10^{-2}$	23.99	0.72	<b>29.05</b>	<b>1.22</b>	8.52	$9.65 \times 10^{-3}$	<b>4.22</b>	$6.53 \times 10^{-3}$
$5 \times 10^6$	0.97	$8.83 \times 10^{-3}$	<b>0.99</b>	$3.85 \times 10^{-3}$	30.26	0.91	<b>33.76</b>	<b>1.03</b>	3.81	$4.69 \times 10^{-3}$	<b>2.42</b>	$3.25 \times 10^{-3}$
$1 \times 10^7$	0.98	$4.31 \times 10^{-3}$	<b>0.99</b>	$2.02 \times 10^{-3}$	33.19	0.83	<b>36.05</b>	<b>0.89</b>	2.68	$3.14 \times 10^{-3}$	<b>1.87</b>	$2.35 \times 10^{-3}$
$1 \times 10^8$	0.99	$4.96 \times 10^{-4}$	<b>0.99</b>	$3.97 \times 10^{-4}$	43.03	0.82	<b>43.96</b>	<b>0.73</b>	0.84	$9.31 \times 10^{-4}$	<b>0.74</b>	$7.36 \times 10^{-4}$
$1 \times 10^9$	0.99	$4.84 \times 10^{-5}$	<b>0.99</b>	$4.64 \times 10^{-5}$	52.97	0.91	<b>53.12</b>	<b>0.89</b>	0.27	$3.26 \times 10^{-4}$	<b>0.26</b>	$3.06 \times 10^{-4}$

<sup>a</sup>Bold indicates the better result for each photon number case under the same metric.

$$\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right), \quad (20)$$

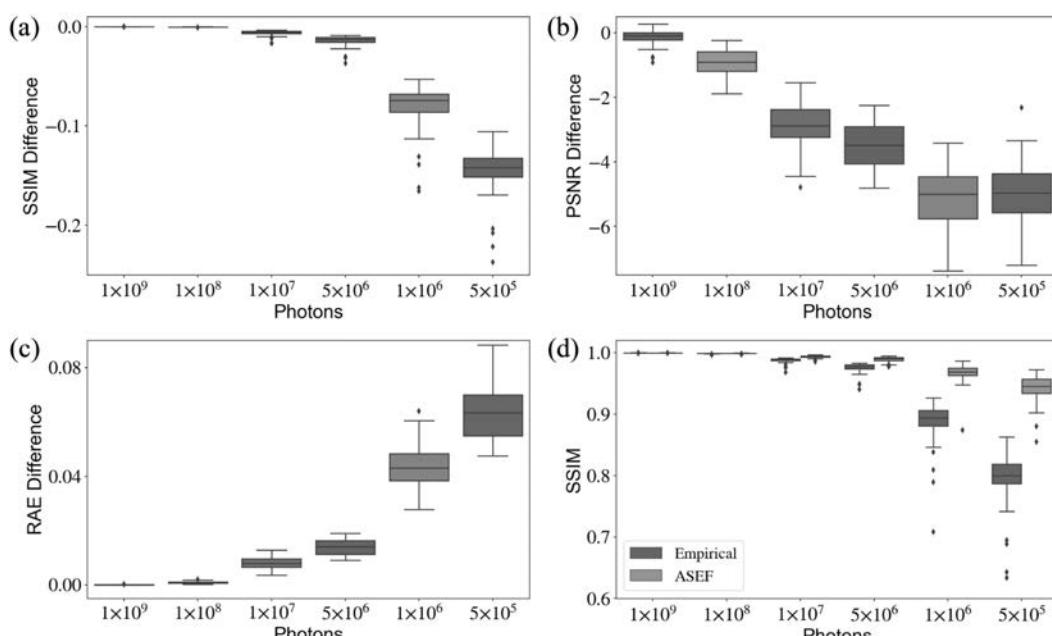
$$\text{RAE} = \sum_{i=1}^m \sum_{j=1}^n \frac{|s(i,j) - s_{gt}(i,j)|}{|s_{gt}(i,j)|}, \quad (21)$$

where  $\text{MSE} = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n [s(i,j) - s_{gt}(i,j)]^2$ . MAX denotes the maximum value of scatter signal  $s$  with dimension  $m \times n$ .

The SSIM, PSNR, and RAE are treated as three metrics to compare the smoothed scatter results using *Empirical* and *ASEF* methods. The result across all testing cases is summarized in Table 3. It is observed that the RAE of *ASEF* is minimal among all different orders of photon numbers, whereas the SSIM and PSNR are maximal compared to the *Empirical* method. Table 3 suggests that for the  $5 \times 10^5$  source photons case referred to as the most challenging case in this study, the *ASEF* can achieve satisfactory performance with SSIM > 0.94, PSNR >

26.55 dB, and RAE < 5.62%. Conversely, the SSIM, PSNR, and RAE of *Empirical* are 0.79, 21.54 dB, and 12.03%, respectively. The boxplots of the metric difference between *Empirical* and *ASEF* are plotted in Figs. 8(a)–8(c). We define metric difference as  $\text{metric}_{\text{diff}} = \text{metric}_{\text{Empirical}} - \text{metric}_{\text{ASEF}}$ . Note that the black line in the box denotes the median of difference statistics for each case. In the figure, the absolute value of the difference statistically becomes larger in the SSIM, PSNR, and RAE metrics with the decrease of the source photons. Figure 8(d) shows the SSIM comparison of *Empirical* and *ASEF* methods, where it is observed that *ASEF* is more robust among different photon number cases.

To demonstrate the efficiency of the proposed framework, we recorded the computation time of the MC particle sampling algorithm and DRL scheme for one scatter image generation using different photon numbers. MC and DRL are speeded up with one Nvidia TITAN Z GPU with 6 GB memory. Table 4 indicates that the computation burden of the MC algorithm exponentially reduces as the number of photons



**Fig. 8.** (a)–(c) indicate boxplots of the metric difference of SSIM, PSNR, and RAE between *Empirical* and *ASEF* for all testing cases.  $\text{metric}_{\text{diff}} = \text{metric}_{\text{Empirical}} - \text{metric}_{\text{ASEF}}$ , where metric denotes SSIM, PSNR, and RAE, respectively. (d) is the boxplot of the SSIM comparison of *Empirical* and *ASEF*.

**Table 4.** Computation Time for One Scatter Image of a Prostate Patient across Different Photon Numbers

	Computation Time (s)							
	$5 \times 10^5$	$1 \times 10^6$	$5 \times 10^6$	$1 \times 10^7$	$1 \times 10^8$	$1 \times 10^9$	$1 \times 10^{10}$	$1 \times 10^{11}$
MC	0.43	0.45	0.57	0.83	5.94	60.00	633.95	6402.60
DRL	8.98	4.80	1.94	0.98	0.32	0.29	0.29	0.29
Total	9.41	5.25	2.51	1.81	6.26	60.29	634.24	6402.89

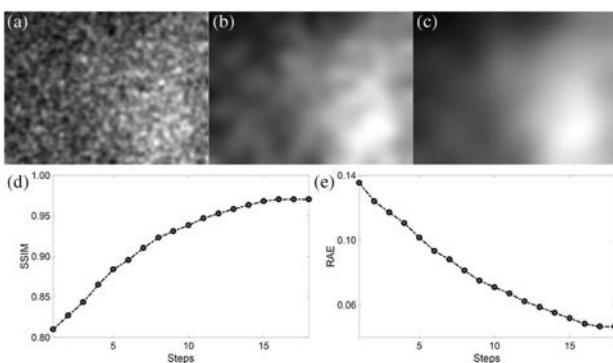
decrease. The DRL scheme is an interaction process of the neural network and the noise removal algorithm. The computation cost of the neural network is negligible, so the computation time of DRL is dominated by the noise removal algorithm. Note that the DRL process is comprised of a sequence of steps, where the raw scatter with fewer photons takes DRL more steps to find out the optimal parameters. To improve the computational efficiency, the DRL framework is initialized with empirical parameters. When the source photons are no less than  $1 \times 10^9$ , DRL just takes one step to reach the optimal solution, so the computation time of DRL for the last three photon numbers in Table 4 is fixed. Briefly, the overall computation time of the proposed ASEF can be reduced to  $\sim 1.81$  s without loss of quality, an outcome which is extremely fast in the MC simulation domain.

#### D. Automatic Scatter Estimation Process

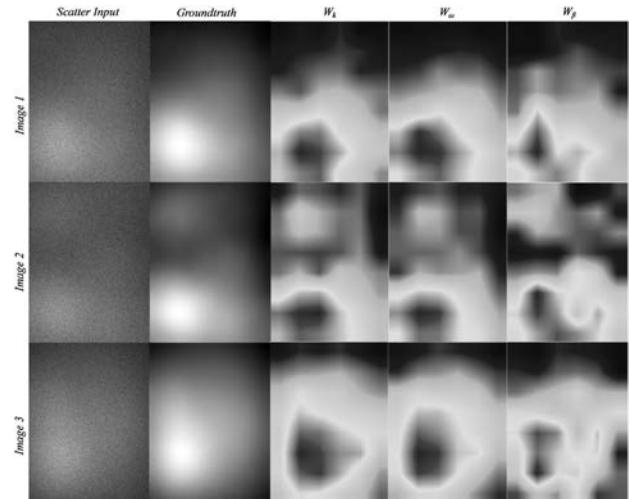
Figure 9 presents the result during the automatic scatter estimation of the ASEF for a testing case. Figures 9(a)–9(c) are smoothed scatter images at Steps 1, 7, and 13, respectively. Visually, the image quality of the scatter signal is gradually ameliorated with a sequence of tuning steps. This outcome implies that the DDQN has learned an optimal policy to wisely make decisions for parameter adjustment. The SSIM and RAE results over steps are shown in Figs. 9(d) and 9(e). As expected, the SSIM increases step by step, whereas the RAE decreases over steps. Thus, the ASEF can improve scatter image quality intelligently.

#### E. Grad-CAM Heatmap

The Grad-CAM heatmaps of the three subnetworks  $\{W_k, W_\omega, W_\beta\}$  are simultaneously produced when ASEF is verified in the testing case. These heatmaps are upsampled to match the original scatter image size for exhibition and nor-



**Fig. 9.** Automatic scatter estimation process for a testing case. (a)–(c) are smoothed scatter images at Steps 1, 7, and 13, respectively. (d) and (e) separately plot the SSIM and RAE over steps.

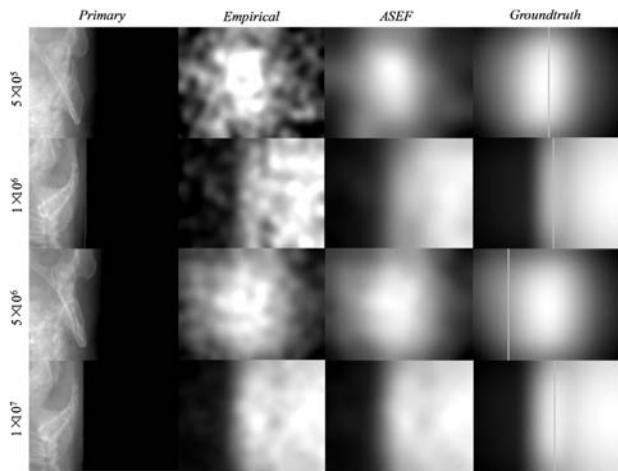


**Fig. 10.** Different scatter images. From left to right: scatter projection input, the ground truth of the scatter image at the first column, and Grad-CAM heatmaps of three subnetworks  $\{W_k, W_\omega, W_\beta\}$ .

malized in [0,1], where the prominent region denoted as the red area makes critical contributions to action choice.  $\{W_k, W_\omega, W_\beta\}$  take the first column of Fig. 10 to predict an action and produce the Grad-CAM heatmaps for each subnetwork shown in the last three columns. The second column is the corresponding ground truth of the first column for visualization comparison. The rows in Fig. 10 represent different scatter images. Note that the high-intensity region of the scatter image in the first row is at the bottom left, whereas the high-intensity area at the third row is close to the image center. The image at the second row has two highlighted regions, for which the major and minor areas are separately located at the bottom and top left. The salient region of the heatmap is approximately located in a high scatter intensity area. The Grad-CAM heatmaps follow a similar scatter intensity distribution of the input state (first column), thereby implying that all three subnetworks consider higher values as crucial features for action prediction. This inference is reasonable because higher scatter intensities are more likely to suffer from noise contamination compared to lower scatter intensity signals.

#### F. Generalization Verification

All of the above training and testing cases are the scatter images of a head and neck (H&N) patient; since the proposed framework can improve image quality dynamically in a manner similar to human intelligence, we utilize scatter images of a prostate patient to further verify it. As depicted by the primary signals in Fig. 11, imaging of the pelvis is so large that it exceeds the field



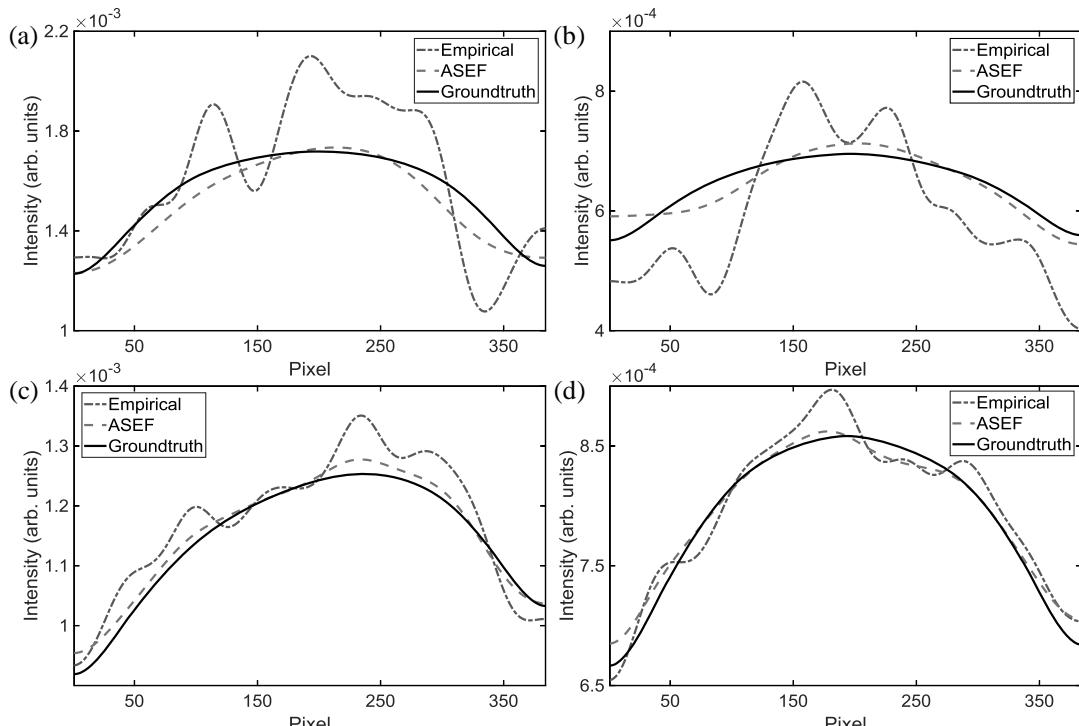
**Fig. 11.** From top to bottom: four prostate cases with  $5 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ , and  $1 \times 10^7$  source photons. From left to right: primary signals, smoothed scatter signals restored by the over-relaxation algorithm with empirical parameters, smoothed scatter signals restored by the proposed framework, and the ground truth.

of view (FOV) of a CBCT scan. Thus, the projections of the prostate structure are usually acquired by means of half-fan scanning protocol. It is observed that *ASEF* is superior compared with *Empirical* for all prostate cases. Moreover, the intensity profiles of *ASEF* in Fig. 12 are more similar to the corresponding lines of ground truth, in comparison to the tortuous *Empirical* lines.

#### 4. DISCUSSION

As verified in Fig. 3 and Table 4, producing high-quality particle distribution only by the MC algorithm is time-consuming, although the process is speeded up using GPUs. Reducing source photons can considerably improve computational efficiency but lead to noise contamination on the scatter signals. Therefore, an efficient noise removal algorithm based on Poisson distribution and sparse feature fusion is proposed to solve the above issue. The suggested algorithm can effectively make the quality of the scatter signal close to the undistorted signal but degrade gradually for a lower photon number (Figs. 4 and 5). High photon number cases (source photons of over  $1 \times 10^9$ ) can be used for experimental study, which requires good quality at the cost of computation time. Meanwhile, ultra-low photon number cases are applied for clinics since the computation time is considered as a crucial factor in clinics. The DRL scheme is employed to boost the performance of the noise removal algorithm for ultra-low photon number cases. In this way, an automatic scatter estimation scheme suitable for the large range of photon numbers is constructed.

In the study, Figs. 6 and 7 depict the proposed framework that can produce satisfactory scatter signals with different source photons. Table 3 summarizes the results of SSIM, PSNR, and RAE across all testing H&N cases and indicates that *ASEF* can achieve good performance with  $\text{SSIM} > 0.94$ ,  $\text{PSNR} > 26.55 \text{ dB}$ , and  $\text{RAE} < 5.62\%$ . Figure 8 further proves that our framework was robust across different photon numbers. The computational time of *ASEF* run on the GPU as recorded in Table 4 reveals that our approach is



**Fig. 12.** (a)–(d) Intensity profiles of the four prostate cases presented in Fig. 11. Profile locations are outlined by orange lines in the last column of Fig. 11.

highly efficient with  $\sim 1.81$  s for one scatter projection generation.

The Grad-CAM heatmap was applied to visualize how the DDQN makes action decisions. The region crucial for action prediction is identified in Fig. 10. Obviously, all three subnetworks had conformal intensity distributions similar to the scatter signal, thereby suggesting that the network tends to focus on high scatter intensity areas in view of the fact that the effect of noise on the higher-value region is greater than that in the lower-value counterpart.

We also validated the generalization of the proposed framework on a prostate patient while the DDQN was trained with scatter images of an H&N patient. The scatter signal of prostate patient is much sparser than the H&N scatter signal because the former is acquired by a half-fan scanning protocol. Four prostate scatter projections with  $5 \times 10^5$ ,  $1 \times 10^6$ ,  $5 \times 10^6$ , and  $1 \times 10^7$  source photons are validated to exhibit the performance. As shown in Fig. 11, the scatter projections by the proposed framework across all cases with different photon numbers were visually superior, and the intensity profiles in Fig. 12 demonstrated the desired performance as well.

With the rapid development of deep-learning algorithms, it is intuitive and straightforward to build end-to-end CNN for scatter removal. Moreover, several researches [36,37] have demonstrated successful applications of this scheme. However, some weaknesses and limitations exist to some extent. Training the CNN requires numerous paired data for which scatter projections calculated by the MC algorithm are generally considered as output labels for supervised learning. As mentioned before, yielding an excellent scatter image with adequate source photons will take substantial time; not to mention, hundreds of scatter projections are required to avoid over-fitting [38]. In addition, scatter signals are highly dependent on the CT geometry configuration and X-ray source energy spectrum [39–42], and the CNN model trained on a dataset may be inappropriate for other types of datasets. Note that only 45 scatter images are used in our DRL training phase. Because of the DRL training strategy illustrated in Table 1, the training datasets stored in the experience replay pool with the capacity of 2000 are  $\{s_t, a_t, r_t, s_{t+1}\}$  rather than 45 images. Furthermore, the training datasets are dynamically updated as the steps change, so there are abundant training data for the neural network model.

## 5. CONCLUSION

In this study, we proposed an ASEF integrating an MC particle simulation algorithm, statistical distribution model fused with sparse feature representation under Poisson distribution, and a DRL scheme for high-efficiency photon distribution estimation. Experimental results demonstrated that our proposed framework has superior performance for the high-efficiency scatter estimation of large range photon numbers.

**Funding.** National Key Research and Development Program of China (2016YFA0202003); National Natural Science Foundation of China (61971463); Ministry of Science and Technology Planning Project of Guangdong

(2015B020233002, 2015B020233005); Guangzhou Science and Technology Plan Project (202002030385).

**Acknowledgment.** We thank Professor Xun Jia and his MC simulation group from the Department of Radiation Oncology at University of Texas Southwestern Medical Center for their assistance with comments and support on this work.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- P. J. Verheijen, "Statistical distributions in particle technology," in *Particle Technology Course* (Department of Chemical Engineering/ Technical University Delft, 2001).
- D. A. Nix and A. S. Weigend, "Estimating the mean and variance of the target probability distribution," in *Proceedings of the IEEE International Conference on Neural Networks (ICNN'94)* (1994), pp. 55–60.
- L. A. Dawson and D. A. Jaffray, "Advances in image-guided radiation therapy," *J. Clin. Oncol.* **25**, 938–946 (2007).
- L. A. Love and R. A. Kruger, "Scatter estimation for a digital radiographic system using convolution filtering," *Med. Phys.* **14**, 178–185 (1987).
- J. Star-Lack, M. Sun, A. Kaestner, R. Hassanein, G. Virshup, T. Berkus, and M. Oelhafen, "Efficient scatter correction using asymmetric kernels," *Proc. SPIE* **7258**, 72581Z (2009).
- R. Ning, X. Tang, and D. Conover, "X-ray scatter correction algorithm for cone beam CT imaging," *Med. Phys.* **31**, 1195–1202 (2004).
- J. S. Maltz, B. Gangadharan, M. Vidal, A. Paidi, S. Bose, B. A. Faddegon, M. Aubin, O. Morin, J. Pouliot, and Z. Zheng, "Focused beam-stop array for the measurement of scatter in megavoltage portal and cone beam CT imaging," *Med. Phys.* **35**, 2452–2462 (2008).
- L. Zhu, N. R. Bennett, and R. Fahrig, "Scatter correction method for X-ray CT using primary modulation: theory and preliminary results," *IEEE Trans. Med. Imaging* **25**, 1573–1587 (2006).
- G. Poludniowski, P. M. Evans, V. N. Hansen, and S. Webb, "An efficient Monte Carlo-based algorithm for scatter correction in keV cone-beam CT," *Phys. Med. Biol.* **54**, 3847–3864 (2009).
- E. Mainegra-Hing and I. Kawrakow, "Fast Monte Carlo calculation of scatter corrections for CBCT images," *J. Phys. Conf. Ser.* **102**, 012017 (2008).
- X. Jia, H. Yan, L. Cervino, M. Folkerts, and S. B. Jiang, "A GPU tool for efficient, accurate, and realistic simulation of cone beam CT projections," *Med. Phys.* **39**, 7368–7378 (2012).
- Y. Xu, Y. Chen, Z. Tian, X. Jia, and L. Zhou, "Metropolis Monte Carlo simulation scheme for fast scattered X-ray photon calculation in CT," *Opt. Express* **27**, 1262–1275 (2019).
- A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems* (2012), pp. 1097–1105.
- J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), pp. 3431–3440.
- O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), pp. 234–241.
- K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
- G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4700–4708.
- W. Zhao, B. Han, Y. Yang, M. Buyounouski, S. L. Hancock, H. Bagshaw, and L. Xing, "Incorporating imaging information from deep

- neural network layers into image guided radiation therapy (IGRT)," *Radiotherapy Oncol.* **140**, 167–174 (2019).
19. V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, and G. Ostrovski, "Human-level control through deep reinforcement learning," *Nature* **518**, 529–533 (2015).
  20. D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, and M. Lanctot, "Mastering the game of Go with deep neural networks and tree search," *Nature* **529**, 484–489 (2016).
  21. C. Shen, Y. Gonzalez, L. Chen, S. B. Jiang, and X. Jia, "Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning," *IEEE Trans. Med. Imaging* **37**, 1430–1439 (2018).
  22. H. H. Tseng, Y. Luo, S. Cui, J. T. Chien, R. K. Ten Haken, and I. El Naqa, "Deep reinforcement learning for automated radiation adaptation in lung cancer," *Med. Phys.* **44**, 6690–6705 (2017).
  23. C. Shen, Y. Gonzalez, P. Klages, N. Qin, H. Jung, L. Chen, D. Nguyen, S. B. Jiang, and X. Jia, "Intelligent inverse treatment planning via deep reinforcement learning, a proof-of-principle study in high dose-rate Brachytherapy for cervical cancer," *Phys. Med. Biol.* **64**, 115013 (2019).
  24. C. Shen, L. Chen, Y. Gonzalez, and X. Jia, "Improving efficiency of training a virtual treatment planner network via knowledge-guided deep reinforcement learning for intelligent automatic treatment planning of radiotherapy," *arXiv:2007.12591* (2020).
  25. C. Shen, D. Nguyen, L. Chen, Y. Gonzalez, R. McBeth, N. Qin, S. B. Jiang, and X. Jia, "Operating a treatment planning system using a deep-reinforcement learning-based virtual treatment planner for prostate cancer intensity-modulated radiation therapy treatment planning," *Med. Phys.* **47**, 2329–2336 (2020).
  26. J. M. Boone and J. A. Seibert, "Accurate method for computer-generating tungsten anode X-ray spectra from 30 to 140 kV," *Med. Phys.* **24**, 1661–1670 (1997).
  27. X. Jia, H. Yan, X. Gu, and S. B. Jiang, "Fast Monte Carlo simulation for patient-specific CT/CBCT imaging dose calculation," *Phys. Med. Biol.* **57**, 577–590 (2012).
  28. E. Woodcock, T. Murphy, P. Hemmings, and S. Longworth, "Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry," in *Conference on Applications of Computing Methods to Reactor Problems* (1965), pp. 557–579.
  29. E. Jonsson, S. C. Huang, and T. Chan, "Total variation regularization in positron emission tomography," *CAM report 9848* (1998).
  30. T. Le, R. Chartrand, and T. J. Asaki, "A variational approach to reconstructing images corrupted by Poisson noise," *J. Math. Imaging Vis.* **27**, 257–263 (2007).
  31. L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D* **60**, 259–268 (1992).
  32. H. Van Hasselt, A. Guez, and D. Silver, "Deep reinforcement learning with double Q-learning," *arXiv:1509.06461* (2015).
  33. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
  34. R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-Cam: visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 618–626.
  35. M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, and M. Isard, "TensorFlow: a system for large-scale machine learning," in *12th Symposium on Operating Systems Design and Implementation* (2016), pp. 265–283.
  36. M. Knaup, J. Kuntz, S. Sawall, and M. Kachelrieß, "Real-time scatter estimation for medical CT using the deep scatter estimation: method and robustness analysis with respect to different anatomies, dose levels, tube voltages, and data truncation," *Med. Phys.* **46**, 238–249 (2019).
  37. D. C. Hansen, G. Landry, F. Kamp, M. Li, C. Belka, K. Parodi, and C. Kurz, "ScatterNet: a convolutional neural network for cone-beam CT intensity correction," *Med. Phys.* **45**, 4916–4926 (2018).
  38. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
  39. B. Ohnesorge, T. Flohr, and K. Klingenberg-Regn, "Efficient object scatter correction algorithm for third and fourth generation CT scanners," *Eur. Radiol.* **9**, 563–569 (1999).
  40. M. Baer and M. Kachelrieß, "Hybrid scatter correction for CT imaging," *Phys. Med. Biol.* **57**, 6849–6867 (2012).
  41. W. Zhao, S. Brunner, K. Niu, S. Schafer, K. Royalty, and G. H. Chen, "Patient-specific scatter correction for flat-panel detector-based cone-beam CT imaging," *Phys. Med. Biol.* **60**, 1339–1365 (2015).
  42. M. Sun and J. Star-Lack, "Improved scatter correction using adaptive scatter kernel superposition," *Phys. Med. Biol.* **55**, 6695–6720 (2010).

# PHOTONICS Research

## Deep compressed imaging via optimized pattern scanning

KANGNING ZHANG, JUNJIE HU, AND WEIJIAN YANG\* 

Department of Electrical and Computer Engineering, University of California, Davis, California 95616, USA

\*Corresponding author: weijyang@ucdavis.edu

Received 7 October 2020; revised 31 December 2020; accepted 13 January 2021; posted 15 January 2021 (Doc. ID 410556); published 1 March 2021

The need for high-speed imaging in applications such as biomedicine, surveillance, and consumer electronics has called for new developments of imaging systems. While the industrial effort continuously pushes the advance of silicon focal plane array image sensors, imaging through a single-pixel detector has gained significant interest thanks to the development of computational algorithms. Here, we present a new imaging modality, deep compressed imaging via optimized-pattern scanning, which can significantly increase the acquisition speed for a single-detector-based imaging system. We project and scan an illumination pattern across the object and collect the sampling signal with a single-pixel detector. We develop an innovative end-to-end optimized auto-encoder, using a deep neural network and compressed sensing algorithm, to optimize the illumination pattern, which allows us to reconstruct faithfully the image from a small number of measurements, with a high frame rate. Compared with the conventional switching-mask-based single-pixel camera and point-scanning imaging systems, our method achieves a much higher imaging speed, while retaining a similar imaging quality. We experimentally validated this imaging modality in the settings of both continuous-wave illumination and pulsed light illumination and showed high-quality image reconstructions with a high compressed sampling rate. This new compressed sensing modality could be widely applied in different imaging systems, enabling new applications that require high imaging speeds. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.410556>

### 1. INTRODUCTION

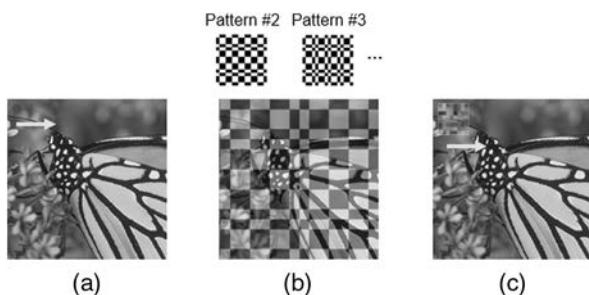
High-speed imaging has become more and more crucial in many new applications, such as in biomedicine, surveillance, and consumer electronics. There are two roadmaps for high-speed optical imaging: engineering a faster focal plane array image sensor and developing new imaging modalities using a single-pixel detector. Although intense industrial efforts have been made, high-speed and low-noise silicon focal plane array cameras are still expensive. Furthermore, imaging at wavelengths outside the silicon sensitivity spectrum can make the focal plane array cameras considerably more complicated [1]. In contrast, imaging through a single-pixel detector, which shrinks a photodetector array down to a single unit [2–9], can enormously reduce the cost and offer additional features such as reduced pixel crosstalk. A popular imaging modality of the single-pixel detector is based on point scanning [5–9] [Fig. 1(a)], for example, the laser-scanning microscopes that are commonly used in biomedicine [5–8]. However, such methods are speed limited due to the point-by-point data acquisition. Another approach with the single-pixel detector relies on compressed sensing (CS) [10–12], represented by the switching-mask-based single-pixel cameras [2–4] [Fig. 1(b)].

There, the images are encoded by a series of spatially well-designed sampling patterns. For each pattern, all pixels across the entire image are summed and collected by the detector. Leveraging the general prior knowledge of sparsity in images, CS is used to reconstruct the image through a small number of measurements [2–4,13,14]. Although the sampling rate can be below the Nyquist criterion, the imaging frame rate is limited by how fast the sampling pattern can be switched and cycled, which is typically conducted by a digital micromirror device (DMD) [15] and operates at <22.7 kHz. Using an LED array to generate the pattern could increase the overall speed, but so far only  $32 \times 32$  pixel images have been demonstrated [16], and it may be expensive to scale up. It does not allow passive light illumination (i.e., structured detection) either.

In this paper, we propose and demonstrate a deep compressed sensing modality, which can significantly increase the imaging speed while preserving a high reconstruction quality. This approach combines the strength of both compressed sensing and point-scanning imaging, and we call it deep compressed imaging via optimized pattern scanning (DeCIOPS). Instead of projecting multiple binary patterns onto the entire object sequentially, we utilize only one gray-scale optimized

pattern and project it to a small subset of the object. We then scan the pattern across the object by using fast scanning mirrors and collect the signal convolutionally using a single-pixel detector [Fig. 1(c)]. Compared with the conventional single-pixel camera, which relies on sequentially switching the sampling pattern on a DMD, our scanning approach significantly increases the sampling speed. Compared with the point-scanning system, our method samples a much larger portion of the object at once and recovers the resolution computationally. This allows a great reduction of the sampling number and thus increases the frame rate. We note that the improvement of imaging speed does not require an increase of light energy. In fact, the required light dosage in our method is smaller than in the conventional point-scanning system due to a reduced number of measurements. We build an auto-encoder framework [17] to optimize the sampling pattern. The image acquisition system is treated as an encoder, where the high-resolution object is encoded through the sampling pattern into a few measurements. We then formulate an iterative shrinkage-thresholding algorithm network (ISTA-Net) [18], a CS-induced neural network inspired by the iterative shrinkage-thresholding algorithm (ISTA) [19], as a decoder to reconstruct the image. This auto-encoder is trained in an end-to-end fashion. Such a framework can learn an optimized sampling pattern and simultaneously recover a high-resolution image by extracting the feature of sparsity and searching the optimal pair of the encoder and decoder with the lowest incoherence [11], which is one key feature of DeCIOPS versus other compressed sensing or deep-learning-based super-resolution imaging modalities [20–23]. This new imaging modality can be accustomed to any light-scanning imaging system and will greatly benefit the high-throughput imaging applications.

This paper is organized as follows. In Section 2, we introduce the mechanisms of the imaging modality in two configurations of illuminations, continuous-wave (CW) and pulsed light source, as well as the auto-encoder framework and the deep compressed sensing neural network for optimizing the imaging and reconstruction. In Sections 3 and 4, we show the simulation results and experimental results. In Section 5, we discuss the system performance under different signal-to-noise



**Fig. 1.** Landscape of imaging methods using a single-pixel detector. (a) Point scanning system where the signal from an individual pixel is sequentially recorded. (b) A conventional single-pixel camera where different patterns are sequentially projected on the entire object, and the overlap integrals between the object and each pattern are measured. (c) Deep compressed imaging via optimized pattern scanning (DeCIOPS), where a pattern is scanned across the object, and the subsampled convolution between the pattern and the object is measured.

ratio (SNR) and compression ratio and how DeCIOPS can be applied in two-photon microscopy and passive lighting conditions.

## 2. PRINCIPLE

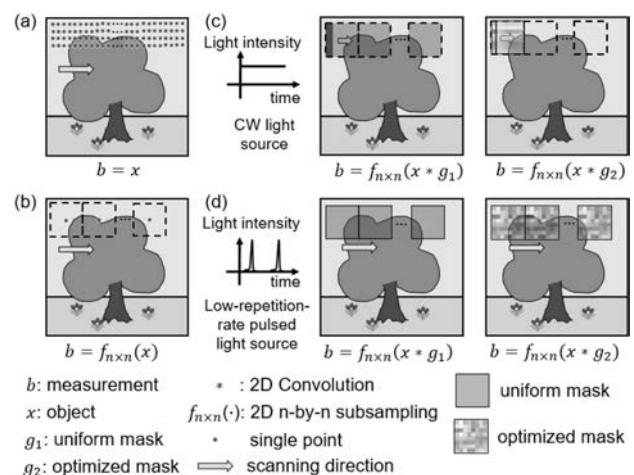
### A. Image Formation

In DeCIOPS, an illumination pattern is generated through a mask and is scanned across the object by using a set of scanning mirrors. The detector records a subsample of the two-dimensional (2D) convolution between the pattern and the object (Fig. 2). Similar to any point-scanning imaging system, DeCIOPS can use either CW or pulsed light sources. The former is commonly used in imaging systems, whereas the latter is specialized for nonlinear microscopy. When a CW light source or high-repetition-rate pulsed light source is used, the detector continuously integrates the signal as the pattern scans. Hence, in DeCIOPS, we project a rectangle-shaped pattern to the object. By finely adjusting the integration time, each acquisition measures a square subset of the object with a desired resolution [Fig. 2(c)]. In case of a low-repetition-rate pulsed light source, the sampling time stamps of the detector are synchronized with the pulse train, and a square-shaped pattern is projected to the object. By matching the spatial sampling step with the size of the pattern, the entire object is sampled appropriately [Fig. 2(d)].

Mathematically, in low-repetition-rate pulsed light illumination, the image formation of DeCIOPS can be expressed as

$$\mathbf{b} = f_{n \times n}(\mathbf{x} * \mathbf{g}), \quad (1)$$

where  $\mathbf{x}$  is the object,  $\mathbf{g}$  is a square-shaped illumination pattern,  $*$  represents the 2D convolution,  $f_{n \times n}(\cdot)$  models the  $n \times n$  undersampling, and  $\mathbf{b}$  is the measured image. Here, we assume that the mask has a size of  $n \times n$  pixels. In the particular case



**Fig. 2.** Schematic of the undersampling schemes in DeCIOPS. (a) Conventional pixel-by-pixel point scanning. (b) Pixel-by-pixel point scanning with a simple undersampling scheme. (c) DeCIOPS in a CW light source configuration with an illumination pattern of a uniform mask (left) or an optimized mask (right). (d) DeCIOPS in a low-repetition-rate pulsed light source configuration with a uniform mask (left) or an optimized mask (right) as an illumination pattern. The mathematic formula below each panel illustrates the process of image formation, where  $\mathbf{g}_1$  and  $\mathbf{g}_2$  are both square shaped.

where  $\mathbf{g}$  is uniform mask  $\mathbf{g}_1$  [Fig. 2(d), left], Eq. (1) is equivalent to a naïve undersampling by unweighted averaging of every  $n \times n$  pixels of the full resolution image [Fig. 2(a)] acquired in single-point scanning. As discussed in Section 2.B,  $\mathbf{g}$  can be optimized to achieve the best image reconstruction performance [Fig. 2(d), right].

In the CW light or high-repetition-rate pulsed light illumination case, where the detector continuously integrates the signal, we configure the illumination pattern in a size of  $n \times 1$  pixels. When this pattern sweeps  $n$  columns, the information of  $n \times n$  pixels is integrated into a single measurement [Fig. 2(c)]. We can use the same mathematical formulation as Eq. (1) to model the image formation, where each column in the  $n \times n$  mask  $\mathbf{g}$  is identical.

### B. End-to-End Optimized Auto-Encoder Framework

In DeCIOPS, we build an auto-encoder framework to simultaneously learn the optimized mask pattern and a neural network for image reconstruction (Fig. 3). The encoder block models the image formation through the following expression:

$$\mathbf{b}^+ = f_{n \times n}(\mathbf{x} * \mathbf{g}) + \epsilon = F(\mathbf{x}) + \epsilon = \Phi\mathbf{x} + \epsilon, \quad (2)$$

where  $F$  is an operator,  $\Phi$  is the linear transfer matrix, both equivalent to the subsampled convolution with mask  $\mathbf{g}$  in Eq. (1),  $\epsilon$  is the additive noise inherent in the imaging system, and  $\mathbf{b}^+$  is the measured mask-encoded image.

The decoder takes  $\mathbf{b}^+$  as the input and aims to reconstruct the original object  $\mathbf{x}$  by solving the following convex optimization problem with a constraint of the sparse representation of  $\mathbf{x}$ :

$$\operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\Phi\mathbf{x} - \mathbf{b}^+\|_2^2 + \lambda \|\Psi(\mathbf{x})\|_1, \quad (3)$$

where  $\Psi(\mathbf{x})$  denotes a transform of  $\mathbf{x}$  into a sparse representation under the basis of  $\Psi$ , and  $\lambda$  is a hyperparameter.

The solution of the problem in Eq. (3) can be initialized by calculating the pseudo-inverse of the encoder from the measurement  $\mathbf{b}^+$ . We then use ISTA [19] to find an optimized solution of  $\mathbf{x}$  by iterating the following two steps:

$$\mathbf{r}^{(k)} = \mathbf{x}^{(k-1)} - \rho \Phi^T (\Phi\mathbf{x}^{(k-1)} - \mathbf{b}^+), \quad (4)$$

$$\mathbf{x}^{(k)} = \operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\mathbf{x} - \mathbf{r}^{(k)}\|_2^2 + \lambda \|\Psi(\mathbf{x})\|_1, \quad (5)$$

where  $k$  denotes the  $k$ th ISTA iteration step, and  $\rho$  is the step size.

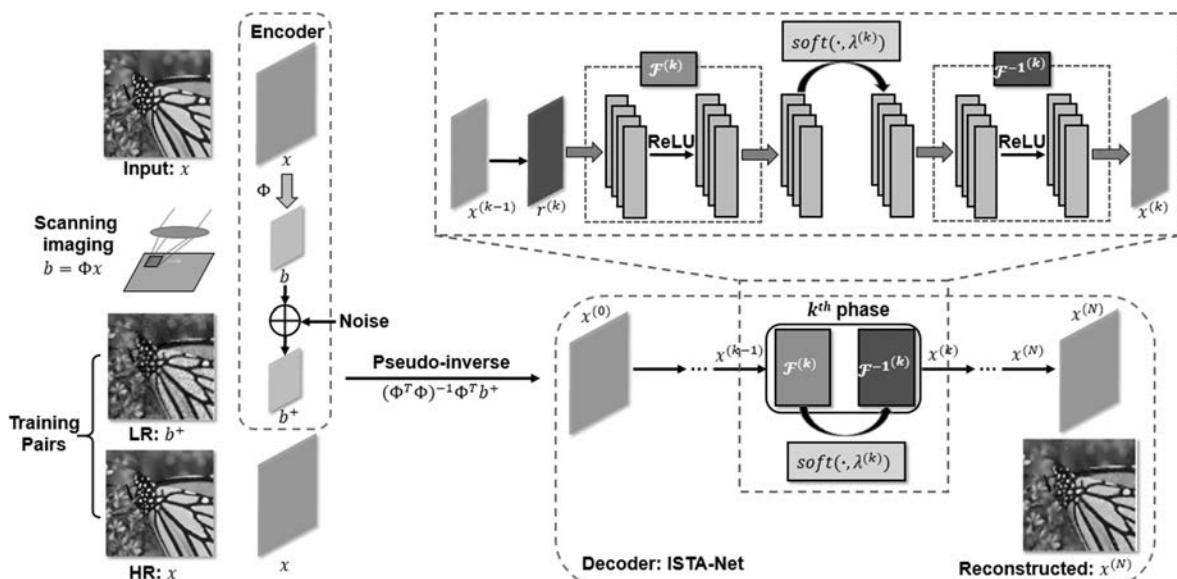
As  $\Psi(\cdot)$  is predefined empirically and may not be suitable for the data, we adopt the ISTA-Net algorithm [18], which can learn  $\Psi(\cdot)$  through the data. In ISTA-Net,  $\Psi(\cdot)$  is replaced with a trainable neural network  $\mathcal{F}(\cdot)$ , and the optimization problem in Eq. (3) can be rephrased into the following L1-norm regularization problem with a nonlinear transform  $\mathcal{F}$ :

$$\operatorname{argmin}_{\mathbf{x}} \frac{1}{2} \|\Phi\mathbf{x} - \mathbf{b}^+\|_2^2 + \lambda \|\mathcal{F}(\mathbf{x})\|_1. \quad (6)$$

The  $k$ th iteration step in the original ISTA is replaced by a series of symmetric learnable parameters in the  $k$ th ISTA-Net phase:

$$\mathcal{F}(\mathbf{x}^{(k)}) = \text{soft}(\mathcal{F}(\mathbf{r}^{(k)}), \theta), \quad (7)$$

$$\mathbf{x}^{(k)} = \mathcal{F}^{-1}(\text{soft}(\mathcal{F}(\mathbf{r}^{(k)}), \theta)), \quad (8)$$



**Fig. 3.** End-to-end optimized auto-encoder framework of image formation and reconstruction in DeCIOPS. The encoder models the image formation. It encodes the high-resolution (HR) object  $\mathbf{x}$  into a low-resolution (LR) output  $\mathbf{b}^+$  through subsampled convolution  $\Phi$  and additive noise. The decoder is implemented with an ISTA-Net, which contains  $N$  phases and reconstructs the object  $\mathbf{x}^{(N)}$ . Each phase is realized by a structure-symmetric pair of a forward transform  $\mathcal{F}^{(k)}$  and a backward transform  $\mathcal{F}^{-1(k)}$  with a soft shrinkage threshold, which factually matches one iteration in the conventional ISTA. ReLU, rectified linear unit; Soft( $\cdot$ ), soft shrinkage threshold.

where  $\theta$  is a learnable parameter in the  $k$ th module,  $\mathcal{F}^{-1}(\cdot)$  is the inverse of  $\mathcal{F}(\cdot)$ , and  $\text{soft}(\cdot)$  represents the soft shrinkage threshold. Finally, we obtain the output of the decoder  $\mathbf{x}^{(N)}$  after a total number of  $N$  ISTA-Net phases.

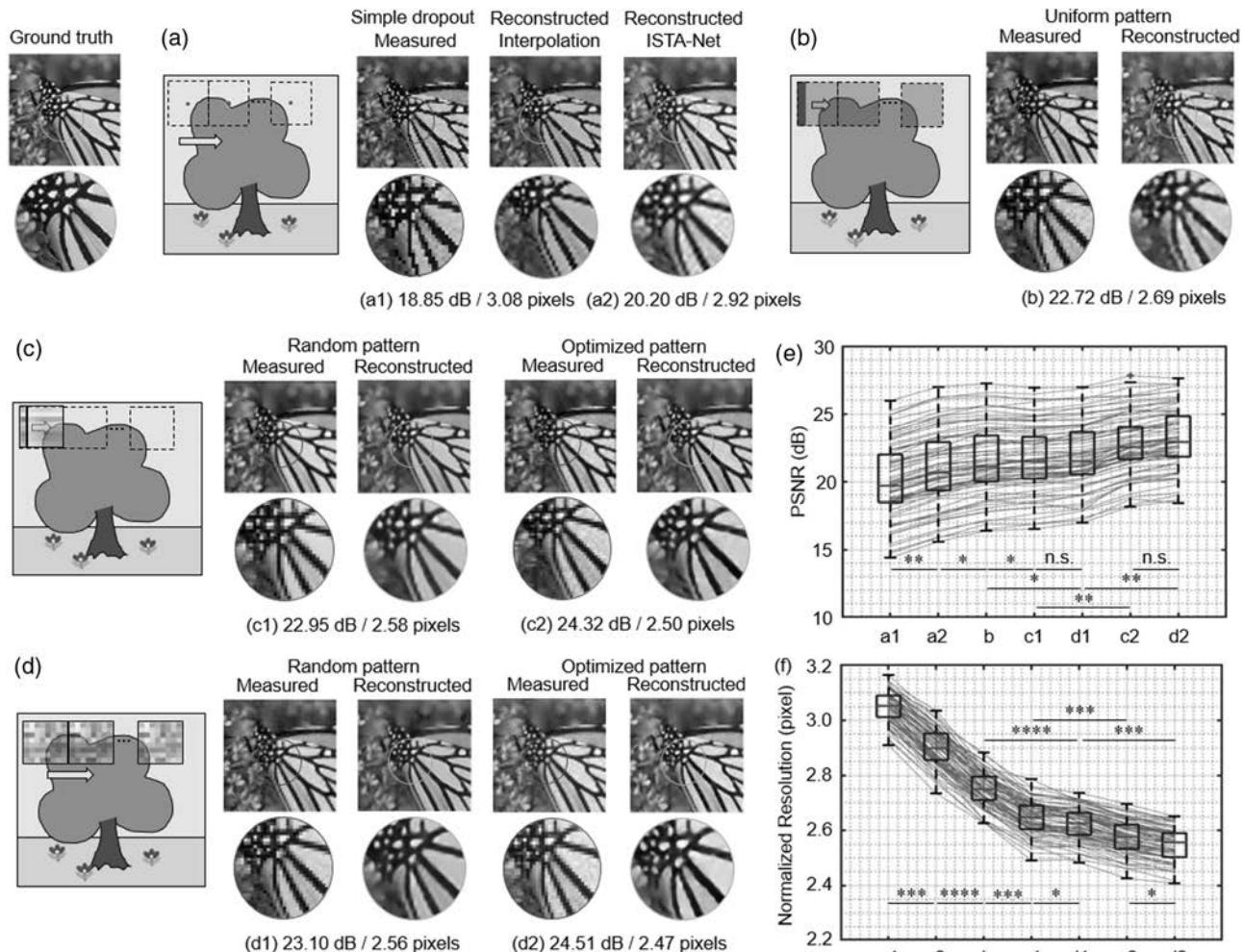
The loss function of ISTA-Net is obtained by calculating the mean square error (MSE) between the output and the ground truth with the constraint of  $\mathcal{F}^{-1} \circ \mathcal{F} = I$ , as both  $\mathcal{F}^{-1}$  and  $\mathcal{F}$  are learnable and symmetric, where  $I$  is the identity operator. As a result, we have the following loss function with the symmetry constraint:

$$\begin{aligned} L_{\text{total}} &= L_{\text{error}} + \gamma L_{\text{constraint}} \\ &= \|\mathbf{x}^{(N)} - \mathbf{x}\|_2^2 - \gamma \left( \sum_{k=1}^N \|\mathcal{F}^{-1(k)}(\mathcal{F}^{(k)}(\mathbf{x})) - \mathbf{x}\|_2^2 \right), \quad (9) \end{aligned}$$

where  $\gamma$  is the weight of the symmetry constraint.

### 3. SIMULATION RESULTS

We trained the auto-encoder using 1500 samples of natural scenes (2D gray-scale image,  $256 \times 256$  pixel size) from ImageNet [24] and validated the model with 79 samples from two widely used benchmark datasets: Set11 [25] and BSD68 [26]. As an illustration, we chose an undersampling rate of 6.25% ( $4 \times 4$  undersampling) and initialized the pattern  $\mathbf{g}$  as a  $4 \times 4$  normalized random Gaussian matrix.  $\gamma$  was set to be 0.01 in the loss function, accompanied by Adam optimization with a learning rate of  $1 \times 10^{-4}$ . We included additive noises in the measurement (5%–10% of the signal strength). The training was performed on a GPU RTX2080Ti 11 GB. The training work of  $N = 9$  ISTA-Net phases takes  $\sim 5$  h for 200 epochs with a batch size of five. We trained two independent auto-encoders, one with a constraint on  $\mathbf{g}$  so each column of  $\mathbf{g}$  is identical (CW light or high-repetition-rate pulsed light illumination) and one without such a constraint (low-repetition-rate



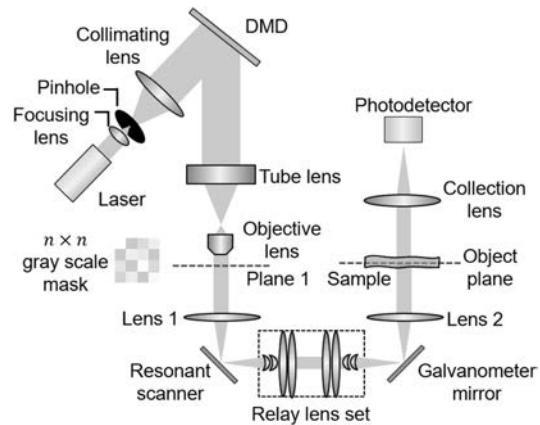
**Fig. 4.** Comparison of the reconstruction performance in the validation data set Set11 and BSD68, at an undersampling rate of 6.25%, through (a) a simple dropout, (b) an unweighted average (uniform pattern), (c) a random or an optimized illumination pattern (DeCIOPS) with a constraint of identical column, and (d) a random or an optimized illumination pattern (DeCIOPS). The PSNR and resolution of the reconstructed images are labeled below the exemplary sample. (e) PSNR of the reconstructed images of all 79 samples in the validation dataset for cases in (a)–(d). (f) Resolution of the reconstructed images of all 79 samples in the validation dataset for cases in (a)–(d). n.s., not significant; \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ , in one-way analysis of variance (ANOVA).

pulsed light illumination). The reconstruction results are evaluated in terms of peak signal-to-noise ratio (PSNR) and spatial resolution by using Fourier ring correlation [27] in the validation dataset. As a control measure, we compared the reconstruction performance of the optimized pattern [Figs. 2(c) and 2(d), right] with a random pattern and two naïve undersampling schemes of the full resolution image either through an unweighted averaging of  $4 \times 4$  pixels [equivalent to the uniform pattern, Figs. 2(c) and 2(d), left] or a simple dropout [i.e., pick one pixel in every  $4 \times 4$  and drop out the others, Fig. 2(b)]. All of the simulation groups employ an independently trained ISTA-Net for image reconstruction. In addition, we also used B-spline interpolation [28] to reconstruct the image that was undersampled through a simple dropout.

We compared the measurements and their corresponding reconstruction results for various schemes (Fig. 4). In all cases, the reconstructed results can resolve higher resolution features than the raw measurement. In the undersampling through simple dropout [Fig. 4(a)], ISTA-Net shows better performance than the interpolation. These results demonstrate the effectiveness of ISTA-Net. As the undersampling through dropout [Fig. 4(a)] misses a substantial amount of information in the original object, its reconstruction result is expected to be the worst. Comparing the undersampling through the uniform pattern [Fig. 4(b)] and the optimized pattern [Figs. 4(c) and 4(d)], we find that the optimized pattern cases show reconstruction results with sharper edges. We reason that this is because the optimized pattern balances both the high-frequency and low-frequency components of the original object during sampling, whereas the uniform pattern performs a low-pass filtering such that the high-frequency component is lost before being reconstructed by the decoder. Indeed, when comparing the quantitative results of the PSNR [Fig. 4(e)] and the resolution of the reconstructed image through Fourier ring correlation [Fig. 4(f)], we find that the optimized pattern case shows the best performance. It is important to note that the optimized pattern also outperforms the exemplary random pattern [Figs. 4(c) and 4(d)] for both PSNR and resolution. This verifies the effectiveness of our end-to-end optimized auto-encoder framework. We also note that the optimized mask without the constraints of identical columns show a better performance in resolution than the one with constraints, though their PSNR does not show a significant difference.

#### 4. EXPERIMENTAL RESULTS

To validate the numerical simulation results, we built an imaging system for DeCIOPS (Fig. 5, more details in Appendix A). Here, we used a DMD to generate the light pattern, as it offers great flexibility in comparing the performance between different patterns. In general, as DeCIOPS requires only one illumination pattern, a fixed pattern mask can be used. To generate a gray-scale mask from the DMD binary pixels, we binned  $32 \times 32$  pixels in the DMD into a super-pixel, which could provide up to 1025 gray-scale levels. A total of  $4 \times 4$  super-pixels were programmed to generate the optimized mask pattern. A 520 nm diode laser source was expanded in beam size and collimated to illuminate the mask pattern on the DMD. The spatially encoded light then passed a  $4f$  system composed of a



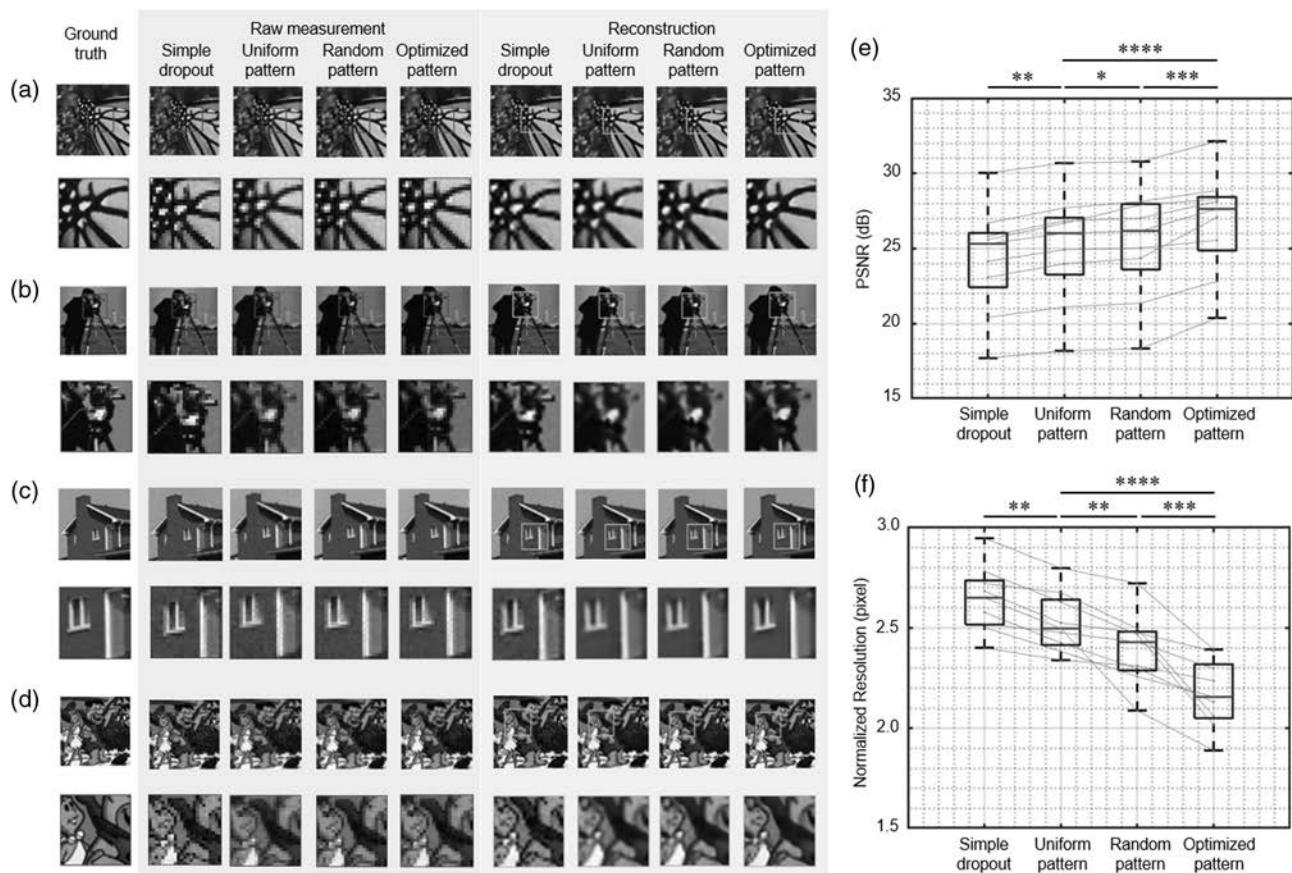
**Fig. 5.** Experiment setup of DeCIOPS. The laser beam is spatially filtered to improve its spatial uniformity and symmetry, collimated and expanded in size, and then incident onto a DMD. The beam is spatially modulated by the DMD and then shrunk in size by a  $4f$  system formed by a tube lens and an objective lens. The light pattern is scanned by a resonant-galvo scanner set, where a resonant scanner and a galvanometer mirror are optically coupled through a relay lens set. The transmitted light from the sample is collected by a photodetector through a collection lens. The  $n \times n$  pattern is generated by the DMD. With an additional  $4f$  system with cylindrical lenses after the objective lens, the  $n \times n$  pattern can be turned into  $n \times 1$  size (Appendix A.1). The red dashed line (plane 1 and object plane) indicates the conjugate plane of the gray-scale pattern mask.

tube lens and an objective lens to reduce its beam size. The light pattern was then raster scanned across the sample through an optically coupled resonant-galvo scanner set. We used a photolithography mask as the sample. The light transmitting through the sample was then collected by a photodetector through a collection lens. While we built this transmission-based imaging system for simplicity, we could turn it into a reflection-based system by adding a beam splitter in front of the sample. We used ScanImage [29] as control software for data acquisition. By adding another  $4f$  system composed of cylindrical lenses right after the objective lens, we could turn the  $4 \times 4$  size pattern into  $4 \times 1$  size (Appendix A.1). Thus, we could use the same setup to validate the different sampling schemes.

As the optical mode from the diode laser did not have uniform intensity, and the pattern could be corrupted by laser interference, we calibrated the DMD to ensure the illumination pattern on the sample plane matched well with the design (Appendix A.3). While our imaging system is naturally a CW light imaging system, we could also mimic the pulsed light source condition through an additional digital sampling step after the image acquisition (Appendix A.5).

#### A. Reconstruction Results with a CW Light Source

We evaluated the experimental results in the natural CW light source setting with the  $4 \times 1$  pattern mask. We compared the reconstruction results across the four undersampling schemes: a simple dropout, the uniform illumination pattern, an exemplary random illumination pattern, and the optimized illumination pattern for various samples (Fig. 6). The ground truth images were obtained by the high-resolution point-scanning



**Fig. 6.** Comparison of the experimental results using different illumination patterns in the scanning in a CW illumination setting. (a)–(d) Experimental results of the sample: (a) butterfly, (b) cameraman, (c) house, and (d) the Flintstones. The different columns show the ground truth results using high-resolution point scanning, raw measurement using different illumination patterns at an undersampling rate of 6.25%, and the corresponding reconstruction results. (e) PSNR of the reconstructed images for a total of nine samples. (f) Spatial resolution of the reconstructed images for a total of nine samples, calculated from Fourier ring correlation. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ , in one-way ANOVA.

approach. At an undersampling rate of 6.25%, all of the reconstructed results through ISTA-Net show a significant improvement from the raw measurement where high-resolution features are better resolved. Comparing between different illumination patterns, we notice more details on the edges of the reconstructed images in the optimized pattern illumination cases [Figs. 6(a)–6(d)]. Using the metric of PSNR [Fig. 6(e)] and spatial resolution [Fig. 6(f)], we find that DeCIOPS outperforms the simple dropout, the uniform illumination pattern, and the random illumination pattern with an average of 2.41 dB, 1.82 dB, and 1.73 dB improvement in PSNR and an average of 18.9%, 14.0%, and 11.6% improvement on spatial resolution, respectively.

## B. Reconstruction Results with a Low-Repetition-Rate Pulsed Light Source

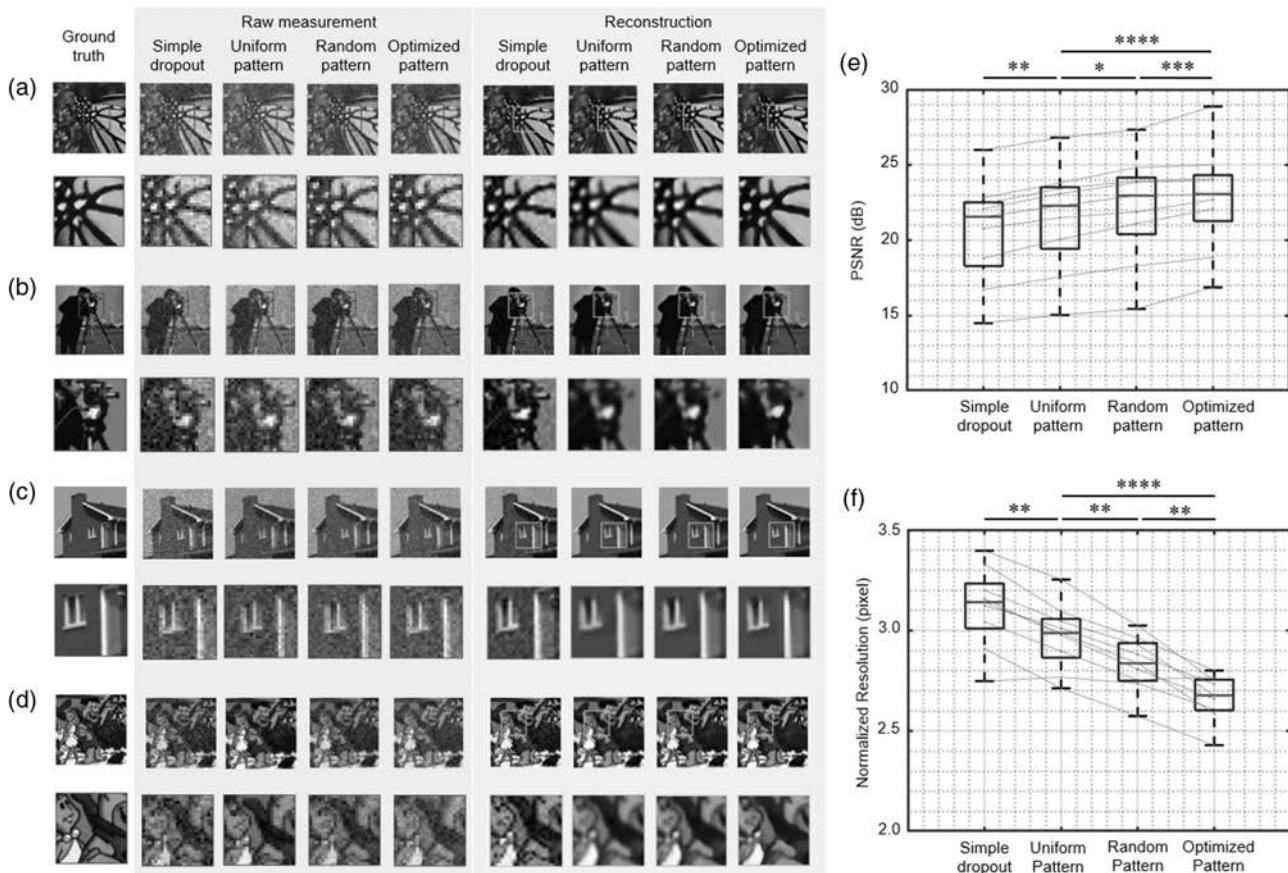
Using the same experimental setup and CW light source, we preprocess the acquired data through a digital undersampling process (Appendix A.5), so the results mimic that acquired by using a low-repetition-rate pulsed light source (Fig. 7). As the integration time of each measured pixel is reduced, the raw measurements have a lower SNR compared to the CW light

source setting. Nevertheless, the reconstruction results show a greatly improved quality. On average, the PSNR of the reconstructed images using the optimized illumination pattern shows an improvement of 1.73 dB, 1.41 dB, and 0.64 dB when they are compared with the simple dropout, uniform illumination pattern, and exemplary random illumination pattern, respectively. The improvement in the spatial resolution in the optimized illumination pattern against the simple dropout, the uniform, and the random illumination pattern is 14.06%, 10.33%, and 7.95%, respectively. Compared with the CW light source setting, the performance improvement of the optimized illumination pattern is reduced because of a reduced SNR in the raw measurement in the low-repetition-rate pulsed light source setting. In the Section 5.A, we further discuss how the SNR influences the reconstructed results.

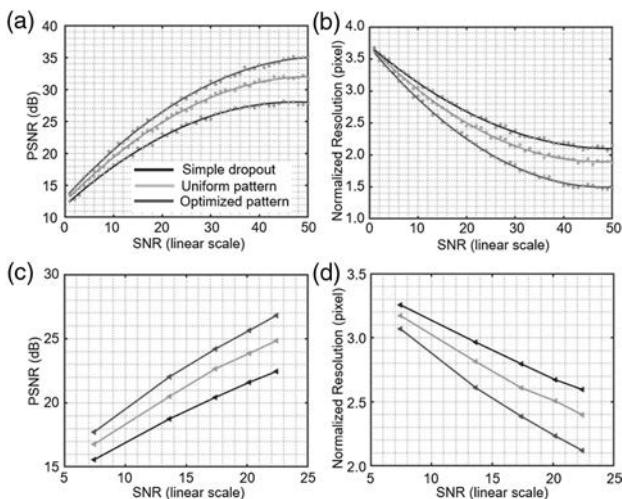
## 5. DISCUSSION

### A. Influence of SNR on Image Reconstruction

In this section, we study the influence of SNR of the raw measurement on the PSNR and pixel resolution of the reconstruction in both simulations and experiments (Fig. 8).



**Fig. 7.** Comparison of the experimental results using different illumination patterns in the scanning in the low-repetition-rate pulsed light illumination setting. (a)–(d) Experimental results of the sample: (a) butterfly, (b) cameraman, (c) house, and (d) the Flintstones. The different columns show the ground truth results using high-resolution point scanning, raw measurement using different illumination patterns at an undersampling rate of 6.25%, and the corresponding reconstruction results. (e) PSNR of the reconstructed images for a total of nine samples. (f) Spatial resolution of the reconstructed images for a total of nine samples, calculated from Fourier ring correlation. \*,  $p < 0.05$ ; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ , in one-way ANOVA.

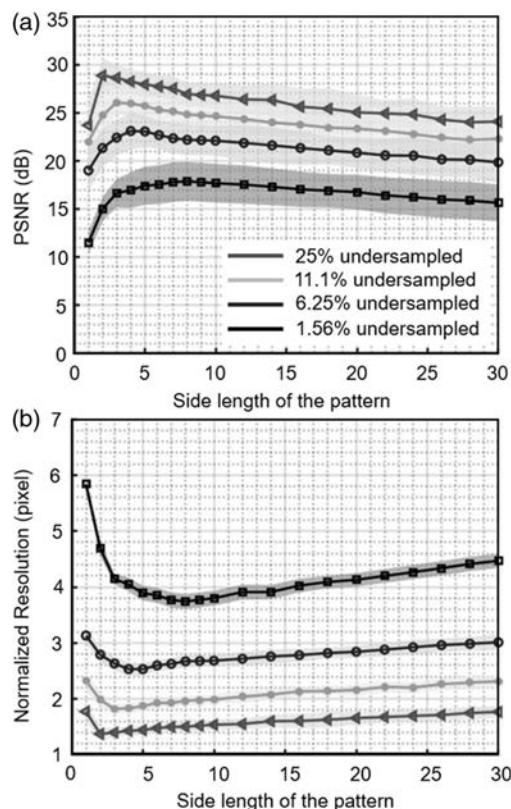


**Fig. 8.** (a) PSNR and (b) pixel resolution of the reconstructed images versus different SNRs in the raw measurement, for three different sampling patterns (CW configuration), performed through simulation at an undersampling rate of 6.25%. The results were averaged across nine samples used in the experiment and fitted with polynomial curves. (c) and (d) show the experimental results averaged across nine samples.

We performed simulations by adding different noise levels in the measurement. We controlled the SNR in experiments by using different light power. For each SNR level, we trained the DeCIOPS independently. As expected, both the PSNR and pixel resolution of the reconstruction improve as the SNR increases and saturates at a high SNR. Compared with the naïve undersampling through simple dropout (blue curve, Fig. 8) and unweighted average (green curve, Fig. 8), the performance of the optimized illumination pattern (red curve, Fig. 8) has a larger improvement in the PSNR and pixel resolution as SNR increases. This phenomenon emphasizes the advantage of the DeCIOPS, where the optimal encoder and decoder are able to match each other better than other independent untrained encoders. As the noise reduces in the raw measurement (i.e., increasing SNR), the network tends to pay more attention on the reconstruction rather than denoising. As a result, the performance difference among different illumination patterns increases.

## B. Compressed Ratio and Size of the Optimized Pattern

In the above simulations and experiments, we set the illumination pattern size such that each pixel in the original



**Fig. 9.** DeCIOPS reconstruction quality (a) PSNR and (b) pixel resolution dependence on the size of the optimized pattern, for an undersampling rate of 25% ( $2 \times 2$ , red), 11.1% ( $3 \times 3$ , green), 6.25% ( $4 \times 4$ , blue), and 1.5625% ( $8 \times 8$ , black), across all 79 samples in the validation dataset. Solid curve, mean; shaded area, standard deviation.

high-resolution object is measured only once during the scanning. We term this as the matching condition between the pattern and undersampling rate. Here, for each specific undersampling rate, we simulated the PSNR of the reconstructed images versus the size of the illumination pattern (Fig. 9). We notice that DeCIOPS generates the best reconstruction results when the size of the pattern matches the 2D undersampling rate, where the imaging modality in the encoder is an algebraic linear orthonormal transform. When the size of the pattern is smaller than that of the matching condition, a sharp drop appears in the reconstruction performance. This is attributed to the permanent information loss from the unsampled pixels in between the adjacent measurements, such that the sensing basis  $\Phi^T$  in Eq. (2) forms a singular transform matrix and thus

degrades the quality of reconstruction. When the size of the pattern is larger than that of the matching condition, the sensing basis  $\Phi^T$  becomes less column orthonormal, resulting in less accurate reconstruction under the CS framework and thus a drop of the PSNR in the reconstructed images.

### C. Optimized Pattern versus Random Pattern

In the conventional switching-mask-based single-pixel camera, the random mask is one of the commonly used sampling bases, as it is incoherent with the spatial property of the sample. We found the same in the pattern-scanning scheme: the random pattern shows a superior performance compared with the uniform pattern. However, the optimized pattern, found by the auto-encoder through the end-to-end optimization, outperforms the random pattern. This is expected, as the random pattern attained by the Monte Carlo method is independent of the dataset of a specific task, unlike the optimized pattern. Furthermore, the performance improvement of the optimized pattern over the random pattern increases when the measurements are more highly undersampled, as seen in the comparison between 6.25% ( $4 \times 4$  pattern) and 1.5625% ( $8 \times 8$  pattern) (Table 1). This is attributed to more trainable parameters in the case of the optimized pattern and more uncertain random variables in the case of the random pattern.

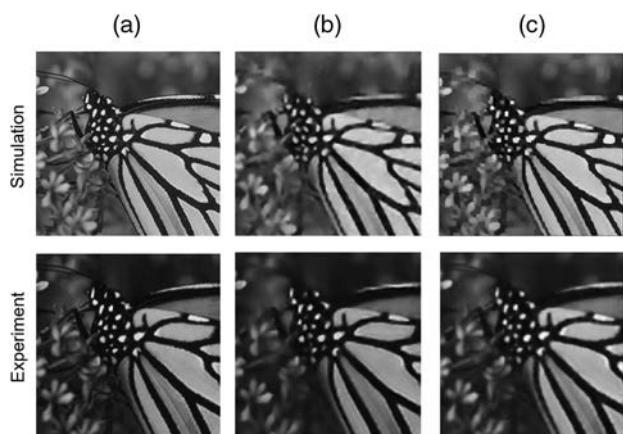
### D. Comparison with Conventional Switching-Mask-Based Single-Pixel Camera

As both DeCIOPS and the conventional switching-mask-based single-pixel camera [Fig. 1(b)] leverage compressed sensing, we expect that their reconstruction quality is similar at the same undersampling rate. Indeed, the one-way analysis of variance (ANOVA) test shows no significant difference for PSNR and pixel resolution in the reconstructed images between these two approaches for both simulation and experiment (Fig. 10, Table 2,  $256 \times 256$  pixels high-resolution objects, 6.25% undersampling rate, ISTA-Net reconstruction framework for both DeCIOPS and switching-mask-based single-pixel camera). However, DeCIOPS has a much faster acquisition speed. To image a high-resolution object with  $256 \times 256$  pixels, DeCIOPS acquires  $64 \times 64$  measurements at an undersampling rate of 6.25%. By using a scanning system with an 8 kHz resonant scanner, the acquisition time is 4 ms. In the conventional switching-mask-based single-pixel camera, at the same undersampling rate, it takes 180 ms to cycle  $64 \times 64 = 4096$  patterns using a 22.7 kHz high-speed DMD, without considering the integration time of the detector. Thus, the imaging speed of our method is orders of magnitude faster than that of the conventional compressed sensing approach using a single-pixel camera.

**Table 1. Comparison of the PSNR and Pixel Resolution across the Uniform Pattern, Random Pattern, and Optimized Pattern between 6.25% and 1.5625% Undersampling Rate in DeCIOPS<sup>a</sup>**

Undersampling Rate	PSNR (dB)			Normalized Resolution (pixel)		
	Uniform Pattern	Random Pattern	Optimized Pattern	Uniform Pattern	Random Pattern	Optimized Pattern
6.25%	$21.49 \pm 1.83$	$21.82 \pm 1.72$	$23.13 \pm 1.54$	$2.79 \pm 0.18$	$2.68 \pm 0.16$	$2.55 \pm 0.15$
1.5625%	$15.91 \pm 2.01$	$16.78 \pm 1.94$	$18.47 \pm 1.75$	$4.55 \pm 0.26$	$4.28 \pm 0.25$	$3.76 \pm 0.21$

<sup>a</sup>CW light configuration, SNR  $\sim 22$ , for all 79 samples in validation dataset.



**Fig. 10.** Comparison of the reconstruction results between DeCIOPS and conventional switching-mask-based single-pixel camera. (a) The ground truth of an original object, butterfly. (b) Reconstruction result of DeCIOPS using ISTA-Net at an undersampling rate of 6.25%. (c) Reconstruction result of the switching-mask-based single-pixel camera imaging approach using ISTA-Net. Top row, simulation. Bottom row, experiment. The ground truth of the experiment is obtained by the high-resolution point scanning

#### E. Advantage of ISTA-Net in Image Reconstruction

In DeCIOPS, we apply ISTA-Net as the decoder. ISTA-Net is a CS-induced neural network. Compared with conventional optimization algorithms where the regularization term is designed empirically, ISTA-Net is entirely data-driven and can learn optimized regularization through the neural network. Compared with other neural networks that could be used for super-resolution, such as U-Net [30,31] and densely connected super-resolution network (DCSRN) [32,33], the embedded CS algorithm in ISTA-Net fits better in the motivation of DeCIOPS and other compressed sensing framework, i.e., performing fewer measurements while being able to reconstruct high-resolution images. Indeed, when we compare the PSNR and pixel resolution of the reconstructed images across ISTA-Net, U-Net, and DCSRN, ISTA-Net shows the best performance (Appendix B). In addition to DeCIOPS, we believe ISTA-Net could also benefit other applications such as denoising [34], fast magnetic resonance imaging (MRI) [35], and other super-resolution imaging modalities [18].

#### F. Advantage of End-to-End Optimized Auto-Encoder and Its Application in Future Imaging Systems

In most existing optical imaging modalities, image formation is empirically designed and optimized, and the deconvolution or

object reconstruction algorithm is subsequently tailored to the image formation process. The recent development of low-cost, advanced micro-optics manufacturing techniques, such as three-dimensional (3D) printing and micro–nano-fabrication [36–40], allows rapid prototyping of user-designed optical elements, which opens new opportunities to redesign the image formation process that best fits the specific applications. Instead of sequentially designing the image formation and the reconstruction algorithm, their joint end-to-end optimization produces a global optimal solution [41–43], which is the underlying principle of DeCIOPS. We use an auto-encoder to model the image formation and reconstruction within a single framework and perform end-to-end training to optimize the sampling pattern and ISTA-Net simultaneously. Our results show that the optimized sampling pattern indeed results in the best overall performance. Such an end-to-end training and data-driven approach prevents any empirical bias that may negatively impact the design. We envision that such an approach will enable many challenging applications such as super-resolution imaging [23,42–46], 3D imaging [41,47–49], and high-speed computational cameras [50–52].

#### G. Applicability in Two-Photon Microscopy

While our imaging system used a CW light source, we mimicked the experimental condition of a pulsed light source and successfully demonstrated the applicability and excellent performance of DeCIOPS. This opens a new avenue to apply DeCIOPS in two-photon microscopy. In conventional two-photon microscopes [5,7,8], the image is acquired through pixel-by-pixel point scanning. While this enables deep tissue imaging as it resists light scattering, it reduces the imaging speed. Recently, there have been multiple reports applying compressed sensing in two-photon microscopy, with the same approach in the conventional switching-mask-based single-pixel camera [53–55]. However, the improvement on the imaging speed is limited due to the low switching speed of DMDs or the liquid-crystal-based spatial light modulators. When applying DeCIOPS in two-photon microscopy, we expect that our approach will significantly increase the imaging speed and will notably benefit functional imaging through two-photon microscopy.

#### H. Passive Light Illumination

In our experiment, DeCIOPS is implemented using active light illumination (i.e., structured illumination), which is commonly used in biomedical imaging. In other imaging systems, passive light illumination may be preferred. In fact, any passive light illumination wide-field imaging using a focal plane array

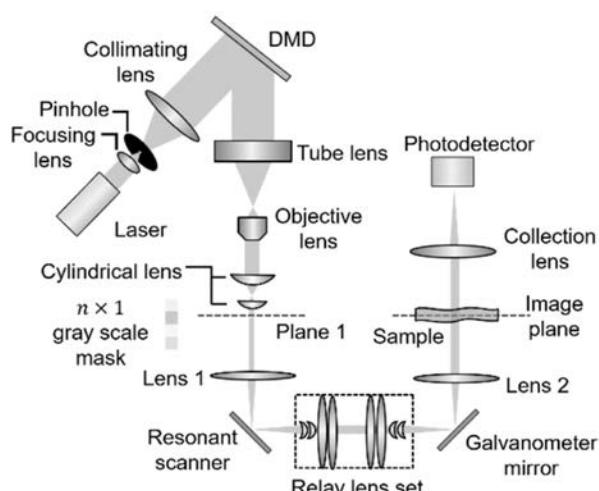
**Table 2. Comparison of Reconstruction Results between DeCIOPS (CW Light Configuration) and Conventional Switching-Mask-Based Single-Pixel Camera<sup>a,b</sup>**

	Simulation		Experiment	
	PSNR (dB)	Normalized Resolution (pixel)	PSNR (dB)	Normalized Resolution (pixel)
DeCIOPS	$28.01 \pm 1.03$	$2.13 \pm 0.12$	$27.71 \pm 1.18$	$2.15 \pm 0.13$
Switching-mask-based single-pixel camera	$28.36 \pm 1.00$	$2.12 \pm 0.12$	$27.98 \pm 1.16$	$2.14 \pm 0.13$

<sup>a</sup>Both simulation and experiment were performed on nine samples from the validation dataset, with an undersampling rate of 6.25% and SNR  $\sim 22$ .

<sup>b</sup>No significant difference (one-way ANOVA) is found on the performance metrics between the two imaging modalities.

(i.e., camera) can be converted to DeCIOPS (Appendix C). A scanner can be added to the passive wide-field imaging to scan the entire image originally projected to the camera. By inserting a fixed mask with an appropriate aperture at the plane where the image is scanned, a single-pixel detector can measure the subsampled convolution between the mask and the original image. The object can then be reconstructed using the same algorithms in DeCIOPS for the active light illumination cases. This way, a structured detection version of DeCIOPS can be implemented.



**Fig. 11.** Experimental setup of DeCIOPS that generates an  $n \times 1$  size pattern and scans it across the sample. The setup is similar to that generating the  $n \times n$  size pattern shown in Fig. 5, but with a  $4f$  system composed of cylindrical lenses added after the objective lens to shrink the original  $n \times n$  size pattern in one dimension by a factor of  $n$  into the  $n \times 1$  size. The red dashed line (plane 1 and object plane) indicates the conjugate plane of the gray-scale pattern mask.

## 6. CONCLUSION

We demonstrated a new high-speed imaging modality, DeCIOPS, by synthesizing the strength of conventional point scanning and single-pixel camera through compressed sensing. The high-speed imaging arises from the fast beam-scanning mechanism and a highly efficient sampling scheme through compressed sensing; meanwhile, an auto-encoder framework allows the simultaneous optimization of the image formation and reconstruction process in DeCIOPS. We validated DeCIOPS through both simulation and experiments in both CW and pulsed light source conditions. This new image modality can be adapted to any existing imaging systems using beam scanning, such as confocal microscope and two-photon microscope, or wide-field cameras with an added scanning system, and will benefit broad applications requiring high-speed imaging.

## APPENDIX A: EXPERIMENTAL SETUP OF DECIOPS

### 1. Optical Setup

The optical setup of DeCIOPS illustrated in Fig. 5 scans an  $n \times n$  pattern on the object and is suitable for low-repetition-rate pulsed light illumination settings. For CW light or high-repetition-rate pulsed light illumination, an  $n \times 1$  pattern is scanned. When this pattern sweeps  $n$  columns, the information of  $n \times n$  pixels can be integrated into a single measurement. To generate the  $n \times 1$  pattern, we set the column to be identical in the  $n \times n$  pattern and add a  $4f$  system composed of cylindrical lenses after the objective lens. The  $n \times n$  pattern is then shrunk in one dimension by a factor of  $n$  into the  $n \times 1$  pattern (Fig. 11). The parameters of the lenses used in the setup are listed in Table 3.

### 2. Pattern Generation

The pattern is generated from a DMD illuminated by a parallel beam. In our experiment, the pattern has  $4 \times 4$  pixels, each of

**Table 3. Detailed Parameters of the Optical Components Used in the Imaging System**

Element	Manufacturer	Part Number	Note
Focusing lens	Thorlabs	A397TM-A	Aspherical lens, focal length 11 mm
Pinhole	Thorlabs	P30D	30 $\mu\text{m}$ pinhole
Collimating lens	Thorlabs	AC254-100-AB-ML	Achromatic lens, focal length 100 mm (for pattern scanning)
Tube lens	Thorlabs	ACL5040-A	Aspherical lens, focal length 40 mm (for point scanning)
Objective lens	Olympus	RMS4X	Focal length 200 mm
Cylindrical lens	Thorlabs	LJ1014L1-A	4 $\times$ objective lens
Lenses 1 and 2	Thorlabs	LJ1227L2-A	Focal length 25.4 mm
Relay lens set [56] (one set listed here; two sets arranged symmetrically are required in setup)	OptoSigma	SLB-60-250 P	Focal length 6.35 mm
	Thorlabs	LC1582-A	Focal length 250 mm
	Thorlabs	LC1582-A	Focal length -75 mm
	Thorlabs	LE1076-A	Focal length -75 mm
	Thorlabs	LA1399-A	Meniscus lens, focal length 100 mm
	Thorlabs	LA1050-A	Focal length 175 mm
	Thorlabs	LA1727-A	Focal length 100 mm
Collection lens	Thorlabs	AC508-075-A-ML	Achromatic lens, focal length 75 mm
Galvanometer scanner	Cambridge Technology	6215HM40B	
Resonant scanner	Cambridge Technology	CRS 8 kHz	
DMD	Texas Instruments	DLP7000	Resonant frequency 8 kHz

which has a gray-scale throughput/intensity ranging from 0 to 1. As each DMD mirror pixel only provides a binary (on or off) light throughput, we group  $32 \times 32$  DMD mirror pixels as one super-pixel, so a gray-scale light throughput becomes feasible within one super-pixel. Each super-pixel thus provides 1025 states in the range from 0 to 1. The multistate light throughput is achieved by randomly selecting parts of the mirrors to be on, the number of which is the product of the desired light throughput and total number of mirror pixels (1024).

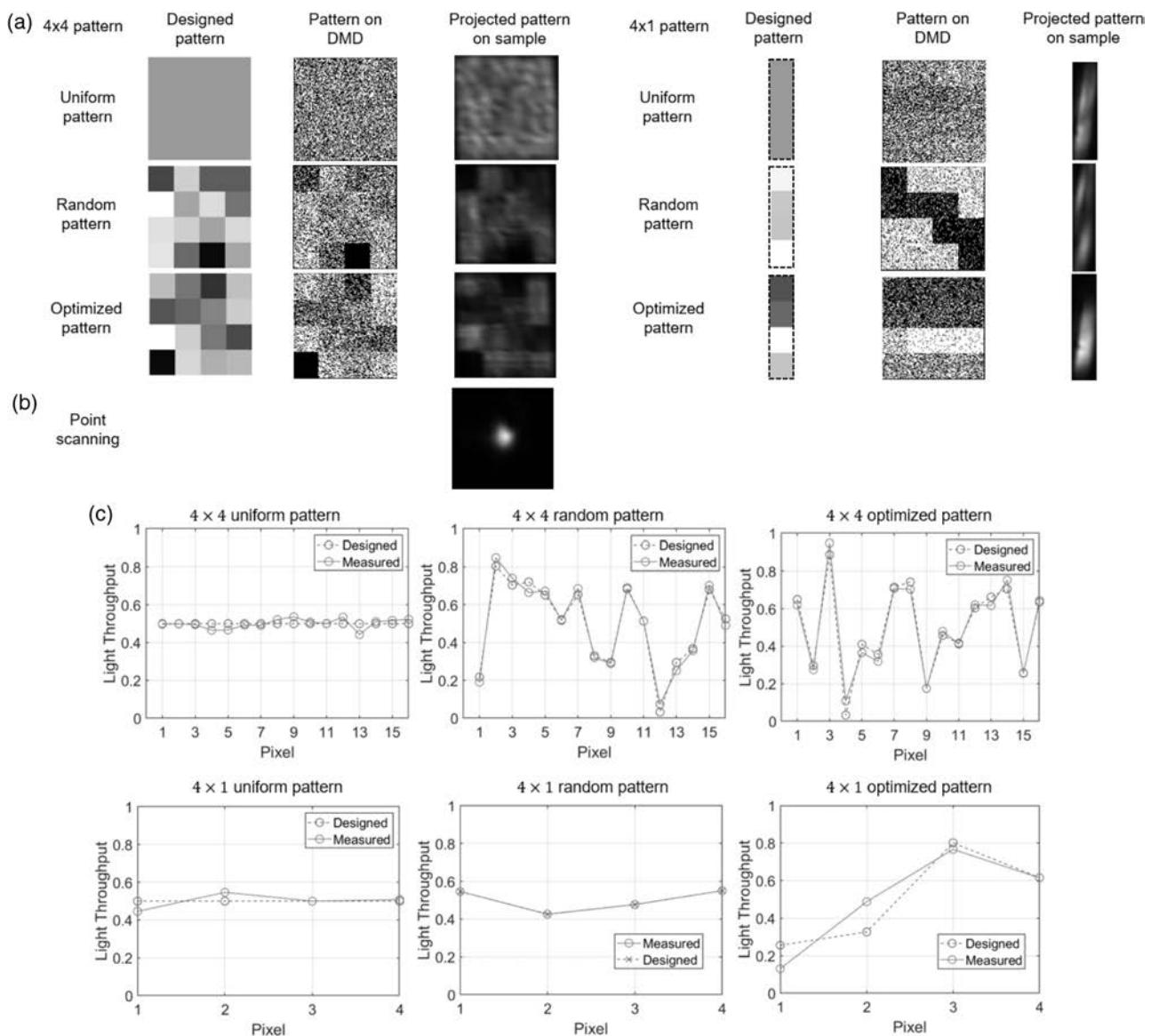
In the case of  $4 \times 1$  pattern scanning, another  $4f$  system composed of cylindrical lenses is added after the objective lens (Fig. 11). Such a system shrinks the  $4 \times 4$  pattern in one dimension by a factor of four.

To obtain the high-resolution ground truth image, we reconfigure the imaging system for point scanning. We replace

the collimating lens between the pinhole and the DMD to the one with shorter focal length and set the displayed pattern on the DMD to be uniform. This reduces the beam size on the image plane, and the pattern effectively turns into a spot matching the size of a single super-pixel of the  $4 \times 4$  pattern. This ensures the same overall energy collected by the photodetector and thus a similar SNR per measurement as in DeCIOPS.

### 3. Characterization of the Pattern

As the optical mode from the diode laser was not symmetric, we implemented a spatial filter using a pinhole after the diode laser output. Nevertheless, non-uniform intensity of the light illuminated on the DMD still appeared. Furthermore, the laser interference could corrupt the pattern. We, thus, fine-tuned the



**Fig. 12.** (a) Measured patterns on the sample (super-pixels) match well with the designed patterns. Each gray-scale super-pixel is generated by  $32 \times 32$  binary pixels in the DMD. The left panel shows the cases for  $4 \times 4$  patterns, and the right panel shows the cases for the  $4 \times 1$  pattern. (b) A single spot pattern is generated for conventional point-scanning imaging to obtain the high-resolution ground truth of the sample. The spot size matches the size of a super-pixel. (c) Pixel-by-pixel comparisons between the measured patterns on the sample and the designed patterns show excellent matchings between the two.

DMD pixel value to calibrate the intensity of the super-pixels, so their values were as close to the design as possible.

We used an iterative approach to calibrate the intensity of the super-pixel. The design pattern  $M_0$  with the gray-scale super-pixel, which was also written as  $M^{(0)}$  for consistency, was first converted to a pattern  $W^{(0)}$  on the DMD through an operator  $\mathcal{D}$ .  $\mathcal{D}$  essentially converted each super-pixel in  $M_0$  into  $32 \times 32$  pixels in binary values on the DMD, with the on and off pixels randomly distributed.  $\mathcal{D}^{-1}$  is the inverse operation. We used a camera to capture the projected pattern  $M^{(1)}$  on the object plane. We then compared  $M^{(1)}$  and  $M_0$  and updated  $W^{(0)}$  into  $W^{(1)}$  through the following algorithm:

$$W^{(1)} = \mathcal{D}((M^{(1)} - M_0) \times a + M^{(0)}), \quad (\text{A1})$$

where  $a$  is the step size to control the update rate. For the  $k$ th iteration, we have

$$W^{(k)} = \mathcal{D}((M^{(k)} - M_0) \times a + \mathcal{D}^{-1}(W^{(k-1)})). \quad (\text{A2})$$

In general, the pattern could be calibrated well with  $<5$  iterations. With this calibration, we verified that the projected pattern on the sample matched the designed pattern in most of the cases (Fig. 12). In the special case of the  $4 \times 1$  random pattern, we took an alternative approach. We randomly generated the pattern and measured it at the object plane. We then used this pattern as the “design” to train the decoder for object reconstruction.

#### 4. Pixel Resolution and Field of View

The size of an individual DMD mirror is  $13.6 \mu\text{m} \times 13.6 \mu\text{m}$ , and thus the size of a super-pixel is  $(32 \times 13.6) \times (32 \times 13.6) = 435.2 \mu\text{m} \times 435.2 \mu\text{m}$ . The imaging system has a magnification of 4.44, resulting in a size of  $\left(\frac{435.2}{4.44}\right) \times \left(\frac{435.2}{4.44}\right) = 97.9 \mu\text{m} \times 97.9 \mu\text{m}$  of a super-pixel on the image plane. A full resolution object with  $256 \times 256$  pixels ( $\sim 24.5 \text{ mm} \times 24.5 \text{ mm}$ ), requires a scanning range of  $\pm 2.8^\circ \times \pm 2.8^\circ$  in the scanner set. We measured the illumination pattern on the object plane across different scanning angles (Fig. 13). The excellent field uniformity ensures the quality of DeCIOPS.

#### 5. Data Acquisition

Data acquisition is performed using a high-speed data acquisition card vDAQ and ScanImage software (Vidrio Technologies). As the illumination pattern continuously scans across each row, the data acquisition card samples the data from the photodetector at a rate higher than the single-pixel rate. The data acquired within the duration of a single pixel is then automatically averaged/integrated and saved as a single-pixel value. Compared with the full resolution single-point-scanning condition ( $256 \times 256$  pixels), we reduce the single-pixel rate and the number of scanning lines by 75% in DeCIOPS (CW light

setting), reaching an undersampling rate of 6.25% ( $64 \times 64$  pixels of measurement).

To mimic the low-repetition-rate pulsed light source condition, we sample the object with a high resolution at  $2048 \times 2048$  pixels (corresponding to a high single-pixel rate) and then downsample the acquired image digitally into  $64 \times 64$  pixels by dropping all of the other pixels. Here, each pixel has a small average/integration duration and could thus be considered as being acquired by a single light pulse.

#### 6. Estimation of Signal-to-Noise Ratio

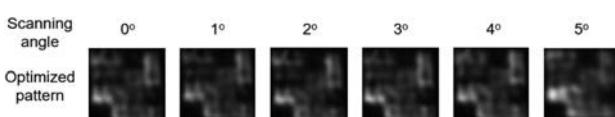
To calculate the SNR of the image in the experiment, we acquire the same image 20 times. For each pixel, we calculate the signal  $\mu$  and the noise  $\sigma$  as the mean and the standard deviation across 20 measurements, respectively. The SNR for the pixel is then estimated as  $\mu/\sigma$ . The SNR of the entire image is taken as the average of the SNR of all pixels.

#### APPENDIX B: COMPARISON BETWEEN ISTA-NET, U-NET, AND DCSRN

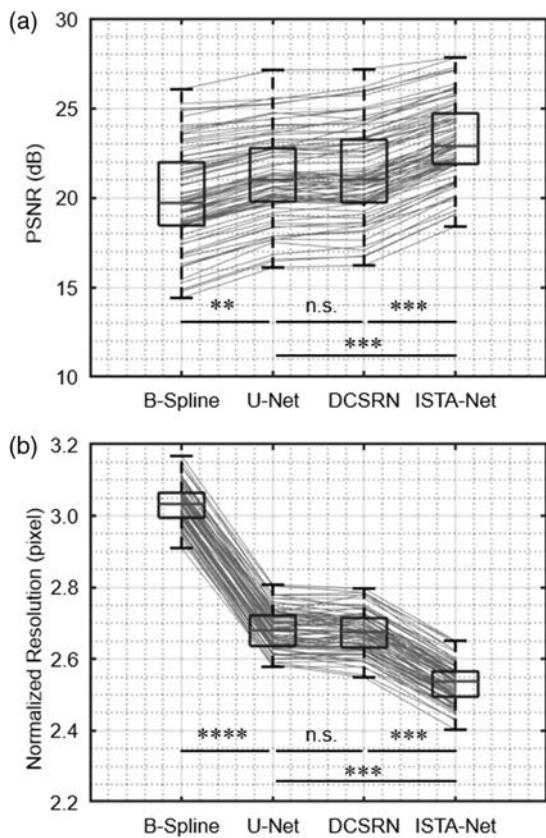
We compare the performance of ISTA-Net, U-Net, and DCSRN in object reconstruction. In the auto-encoder framework, the decoder implemented by ISTA-Net is replaced by U-Net or DCSRN. U-Net is widely used in image reconstruction and segmentation. It first condenses the size of the input images to extract its context and feature and then grows them in an expanding path to perform local reconstruction [30]. DCSRN, derived from densely connected convolutional networks [33], has a faster training speed and accurate reconstruction results and is commonly used in applications such as 2D or 3D biomedical super-resolution imaging. In addition to ISTA-Net, U-Net, and DCSRN, we used B-spline interpolation [28] to reconstruct the object undersampled through a simple dropout as a baseline. All of the simulation was completed on GPU RTX1080Ti 11 GB with 200 epochs and a batch size of five. In each decoder except for the B-spline, we learned an optimized illumination pattern. We used the validation data set to evaluate the PSNR and pixel resolution of the reconstructed objects. U-Net, DCSRN, and ISTA-Net all outperform B-spline interpolation. While U-Net and DCSRN do not show a significant difference in performance, ISTA-Net outperforms both U-Net and DCSRN with a  $\sim 1.85$  dB increase in PSNR and 6.55% improvement in resolution at an undersampling rate of 6.25% (Fig. 14). The simulation results demonstrated a clear advantage of ISTA-Net, which is a CS inspired neural network, in DeCIOPS.

#### APPENDIX C: DECIOPS USING PASSIVE LIGHT ILLUMINATION

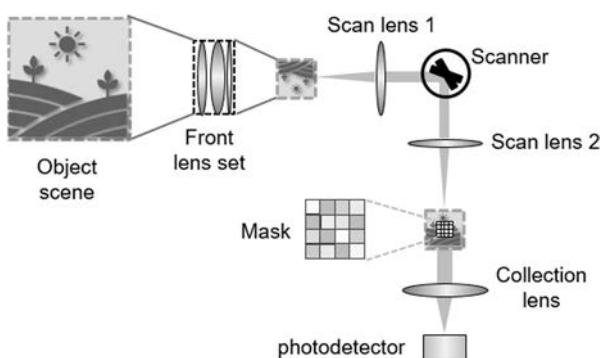
To implement DeCIOPS in a passive light illumination setting (i.e., structured detection), which is commonly used in photography, a scanner can be added to the passive wide-field imaging system to scan the entire image originally projected to the camera. By inserting a fixed mask with an appropriate aperture at the plane where the image is scanned, a single-pixel detector can measure the subsampled convolution between the mask and the original image (Fig. 15). The object can then be



**Fig. 13.** Measured patterns at the image plane stay consistent across different scanning angles.



**Fig. 14.** Comparison of (a) PSNR and (b) pixel resolution of the reconstructed objects of all 79 samples in the validation dataset for B-spline, U-Net, DCSRN, and ISTA-Net in the auto-encoder framework, at an undersampling rate of 6.25%. n.s., not significant; \*\*,  $p < 0.01$ ; \*\*\*,  $p < 0.001$ ; \*\*\*\*,  $p < 0.0001$ , in one-way ANOVA.



**Fig. 15.** Optical setup of DeCIOPS with passive light illumination (i.e., structured detection) for applications such as photography.

reconstructed using the same algorithms in DeCIOPS for the active light illumination cases.

**Funding.** National Science Foundation (CARRER 1847141); National Institute of Neurological Disorders and Stroke (R01NS118289); National Eye Institute (R01NS118289); Burroughs Wellcome Fund (CASI 1015761).

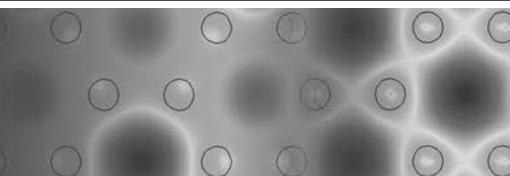
**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- A. Rogalski, *Infrared Detectors* (CRC Press, 2010).
- M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**, 83–91 (2008).
- M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," *Nat. Photonics* **13**, 13–20 (2019).
- G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," *Opt. Express* **28**, 28190–28208 (2020).
- W. Denk, J. H. Strickler, and W. W. Webb, "Two-photon laser scanning fluorescence microscopy," *Science* **248**, 73–76 (1990).
- J. B. Pawley, *Handbook of Biological Confocal Microscopy* (Springer, 2006).
- W. R. Zipfel, R. M. Williams, and W. W. Webb, "Nonlinear magic: multi-photon microscopy in the biosciences," *Nat. Biotechnol.* **21**, 1369–1377 (2003).
- F. Helmchen and W. Denk, "Deep tissue two-photon microscopy," *Nat. Methods* **2**, 932–940 (2005).
- L. Beiser, "Fundamental architecture of optical scanning systems," *Appl. Opt.* **34**, 7307–7317 (1995).
- E. J. Candès, "Compressive sampling," in *Proceedings of the International Congress of Mathematicians* (2006), pp. 1433–1452.
- E. Candes and J. Romberg, "Sparsity and incoherence in compressive sampling," *Inverse Prob.* **23**, 969 (2007).
- E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE Signal Process. Mag.* **25**, 21–30 (2008).
- A. P. Spencer, B. Spokoyny, S. Ray, F. Sarvari, and E. Harel, "Mapping multidimensional electronic structure and ultrafast dynamics with single-element detection and compressive sensing," *Nat. Commun.* **7**, 10434 (2016).
- L. Bian, J. Suo, Q. Dai, and F. Chen, "Experimental comparison of single-pixel imaging algorithms," *J. Opt. Soc. Am. A* **35**, 78–87 (2018).
- J. B. Sampsell, "Digital micromirror device and its application to projection displays," *J. Vac. Sci. Technol. B* **12**, 3242–3246 (1994).
- Z. H. Xu, W. Chen, J. Puenuelas, M. Padgett, and M. J. Sun, "1000 fps computational ghost imaging using LED-based structured illumination," *Opt. Express* **26**, 2427–2434 (2018).
- H. Wu, Z. Zheng, Y. Li, W. Dai, and H. Xiong, "Compressed sensing via a deep convolutional auto-encoder," in *IEEE Visual Communications and Image Processing (VCIP)* (2018), pp. 1–4.
- J. Zhang and B. Ghanem, "ISTA-Net: interpretable optimization-inspired deep network for image compressive sensing," in *IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 1828–1837.
- A. Beck and M. Teboulle, "A fast iterative shrinkage-thresholding algorithm for linear inverse problems," *SIAM J. Imag. Sci.* **2**, 183–202 (2009).
- N. Pavillon and N. I. Smith, "Compressed sensing laser scanning microscopy," *Opt. Express* **24**, 30038–30052 (2016).
- L. Fang, F. Monroe, S. W. Novak, L. Kirk, C. R. Schiavon, S. B. Yu, T. Zhang, M. Wu, K. Kastner, Y. Kubota, Z. Zhang, G. Pekkurnaz, J. Mendenhall, K. Harris, J. Howard, and U. Manor, "Deep learning-based point-scanning super-resolution imaging," *bioRxiv*: 740548 (2019).
- B. Li, Q.-R. Yan, Y.-F. Wang, Y.-B. Yang, and Y.-H. Wang, "A binary sampling Res2net reconstruction network for single-pixel imaging," *Rev. Sci. Instrum.* **91**, 033709 (2020).
- H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Gunaydin, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nat. Methods* **16**, 103–110 (2019).
- J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: a large-scale hierarchical image database," in *IEEE Conference on Computer Vision and Pattern Recognition* (2009), pp. 248–255.
- K. Kulkarni, S. Lohit, P. Turaga, R. Kerviche, and A. Ashok, "Reconnect: non-iterative reconstruction of images from compressively sensed measurements," in *IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 449–458.

26. D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *8th IEEE International Conference on Computer Vision (ICCV)* (2001), pp. 416–423.
27. S. Koho, G. Tortarolo, M. Castello, T. Deguchi, A. Diaspro, and G. Vicidomini, "Fourier ring correlation simplifies image restoration in fluorescence microscopy," *Nat. Commun.* **10**, 3103 (2019).
28. T. M. Lehmann, C. Gonner, and K. Spitzer, "Addendum: B-spline interpolation in medical image processing," *IEEE Trans. Med. Imaging* **20**, 660–665 (2001).
29. T. A. Pologruto, B. L. Sabatini, and K. Svoboda, "ScanImage: flexible software for operating laser scanning microscopes," *Biomed. Eng. Online* **2**, 13 (2003).
30. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), pp. 234–241.
31. J. Feng, J. Deng, Z. Li, Z. Sun, H. Dou, and K. Jia, "End-to-end ResNet based reconstruction algorithm for photoacoustic imaging," *Biomed. Opt. Express* **11**, 5321–5340 (2020).
32. Y. Chen, Y. Xie, Z. Zhou, F. Shi, A. G. Christodoulou, and D. Li, "Brain MRI super resolution using 3D deep densely connected neural networks," in *IEEE 15th International Symposium on Biomedical Imaging (ISBI)* (2018), pp. 739–742.
33. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4700–4708.
34. B. Lee, B. Ku, W.-J. Kim, S. Kim, and H. Ko, "Denoising ISTA-Net: learning based compressive sensing with reinforced non-linearity for side scan sonar image denoising," *J. Acoust. Soc. Korea* **39**, 246–254 (2020).
35. E. D. W. N. Pezzotti, S. Yousefi, M. S. Elmahdy, J. van Gemert, C. Schülke, M. Doneva, T. Nielsen, S. Kastryulin, B. P. F. Lelieveldt, M. J. P. van Osch, and M. Staring, "Adaptive-CS-Net: FastMRI with adaptive intelligence," arXiv:1912.12259 (2019).
36. X. Q. Zhou, Y. H. Hou, and J. Q. Lin, "A review on the processing accuracy of two-photon polymerization," *AIP Adv.* **5**, 030701 (2015).
37. T. Gissibl, S. Thiele, A. Herkommmer, and H. Giessen, "Two-photon direct laser writing of ultracompact multi-lens objectives," *Nat. Photonics* **10**, 554–560 (2016).
38. J. W. Li, P. Fejes, D. Lorenser, B. C. Quirk, P. B. Noble, R. W. Kirk, A. Orth, F. M. Wood, B. C. Gibson, D. D. Sampson, and R. A. McLaughlin, "Two-photon polymerisation 3D printed freeform micro-optics for optical coherence tomography fibre probes," *Sci. Rep.* **8**, 14789 (2018).
39. N. Mohammad, M. Meem, X. Wan, and R. Menon, "Full-color, large area, transmissive holograms enabled by multi-level diffractive optics," *Sci. Rep.* **7**, 5789 (2017).
40. N. Vaidya and O. Solgaard, "3D printed optics with nanometer scale surface roughness," *Microsys. Nanoeng.* **4**, 18 (2018).
41. Y. C. Wu, V. Boominathan, H. J. Chen, A. Sankaranarayanan, and A. Veeraraghavan, "PhaseCam3D-learning phase masks for passive single view depth estimation," in *IEEE International Conference on Computational Photography* (2019), pp. 1–12.
42. V. Sitzmann, S. Diamond, Y. F. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," *ACM Trans. Graph.* **37**, 114 (2018).
43. Q. L. Sun, J. Zhang, X. Dun, B. Ghanem, Y. F. Peng, and W. Heidrich, "End-to-end learned, optically coded super-resolution SPAD camera," *ACM Trans. Graph.* **39**, 9 (2020).
44. Y. Li, B. Sixou, and F. Peyrin, "A review of the deep learning methods for medical images super resolution problems," *IRBM* (2020).
45. W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," *Nat. Biotechnol.* **36**, 460–468 (2018).
46. M. Weigert, U. Schmidt, T. Boothe, A. Muller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, and E. W. Myers, "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nat. Methods* **15**, 1090–1097 (2018).
47. S. Zhang, "High-speed 3D shape measurement with structured light methods: a review," *Opt. Laser Eng.* **106**, 119–131 (2018).
48. Y. C. Wu, Y. Rivenson, H. D. Wang, Y. L. Luo, E. Ben-David, L. A. Bentolila, C. Pritz, and A. Ozcan, "Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning," *Nat. Methods* **16**, 1323–1331 (2019).
49. H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, and L. Waller, "Deep learning for single-shot autofocus microscopy," *Optica* **6**, 794–797 (2019).
50. Y. Ma, Y. Lee, C. Best-Popescu, and L. Gao, "High-speed compressed-sensing fluorescence lifetime imaging microscopy of live cells," *Proc. Natl. Acad. Sci. USA* **118**, e2004176118 (2020).
51. D. Qi, S. Zhang, C. Yang, Y. He, F. Cao, J. Yao, P. Ding, L. Gao, T. Jia, and J. Liang, "Single-shot compressed ultrafast photography: a review," *Adv. Photon.* **2**, 014003 (2020).
52. L. Gao, J. Y. Liang, C. Y. Li, and L. H. V. Wang, "Single-shot compressed ultrafast photography at one hundred billion frames per second," *Nature* **516**, 74–77 (2014).
53. M. Alemdhammad, J. Shin, D. N. Tran, J. R. Stroud, S. P. Chin, T. D. Tran, and M. A. Foster, "Widefield compressive multiphoton microscopy," *Opt. Lett.* **43**, 2989–2992 (2018).
54. P. Wijesinghe, A. Escobet-Montalban, M. Z. Chen, P. R. T. Munro, and K. Dholakia, "Optimal compressive multiphoton imaging at depth using single-pixel detection," *Opt. Lett.* **44**, 4981–4984 (2019).
55. Y.-X. Ren, C. Kong, H. He, X. Zeng, K. K. Tsia, and K. K. Wong, "Encrypted wide-field two-photon microscopy with single-pixel detection and compressed sensing," *Appl. Phys. Express* **13**, 032007 (2020).
56. J. N. Stirman, I. T. Smith, M. W. Kudenov, and S. L. Smith, "Wide field-of-view, multi-region, two-photon imaging of neuronal activity in the mammalian brain," *Nat. Biotechnol.* **34**, 857–862 (2016).

# PHOTONICS Research



## Backpropagation through nonlinear units for the all-optical training of neural networks

XIANXIN GUO,<sup>1,2,3,5,†</sup> THOMAS D. BARRETT,<sup>2,6,†</sup> ZHIMING M. WANG,<sup>1,7</sup> AND A. I. LVOVSKY<sup>2,4,8</sup>

<sup>1</sup>Institute of Fundamental and Frontier Sciences, University of Electronic Science and Technology of China, Chengdu 610054, China

<sup>2</sup>Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK

<sup>3</sup>Institute for Quantum Science and Technology, University of Calgary, Calgary, Alberta T2N 1N4, Canada

<sup>4</sup>Russian Quantum Center, Skolkovo 143025, Moscow, Russia

<sup>5</sup>e-mail: xianxin.guo@physics.ox.ac.uk

<sup>6</sup>e-mail: thomas.barrett@physics.ox.ac.uk

<sup>7</sup>e-mail: zhmwang@uestc.edu.cn

<sup>8</sup>e-mail: alex.lvovsky@physics.ox.ac.uk

Received 8 October 2020; revised 9 January 2021; accepted 11 January 2021; posted 12 January 2021 (Doc. ID 411104); published 1 March 2021

We propose a practical scheme for end-to-end optical backpropagation in neural networks. Using saturable absorption for the nonlinear units, we find that the backward-propagating gradients required to train the network can be approximated in a surprisingly simple pump-probe scheme that requires only simple passive optical elements. Simulations show that, with readily obtainable optical depths, our approach can achieve equivalent performance to state-of-the-art computational networks on image classification benchmarks, even in deep networks with multiple sequential gradient approximation. With backpropagation through nonlinear units being an outstanding challenge to the field, this work provides a feasible path toward truly all-optical neural networks. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.411104>

### 1. INTRODUCTION

Machine learning (ML) is changing the way in which we approach complex tasks, with applications ranging from natural language processing [1] and image recognition [2] to artificial intelligence [3] and fundamental science [4,5]. At the heart (or “brain”) of this revolution are artificial neural networks (ANNs), which are universal function approximators [6,7] capable, in principle, of representing an arbitrary mapping of inputs to outputs. Remarkably, their function only requires two basic operations: matrix multiplication to communicate information between layers, and some nonlinear transformation of individual neuron states (activation function). The former accounts for most of the computational cost associated with ML. This operation can, however, be readily implemented by leveraging the coherence and superposition properties of linear optics [8]. Optics is therefore an attractive platform for realizing the next generation of neural networks, promising faster computation with low power consumption [9–13].

Proposals for optical neural networks (ONNs) have been around for over 30 years [14,15] and have been realized in both free-space [16–18] and integrated [9,10] settings. However, the true power of neural networks is not only that they can approximate arbitrary functions, but also that they can “learn” that approximation. The training of neural networks is, almost

universally, achieved by the backpropagation algorithm [19]. Implementing this algorithm optically is challenging because it requires the response of the network’s nonlinear elements to be different for light propagating forward or backward. Confronted with these challenges, existing ONNs are actually trained with, or heavily aided by, digital computers [9,16,18,20]. As a result, the great advantages offered by optics remain largely unexploited. Developing an all-optically trained ONN to leverage these advantages remains an unsolved problem. Here, we address this challenge and present a practical training method capable of backpropagating the error signal through nonlinear neurons in a single optical pass.

The backpropagation algorithm aims to minimize a loss function that quantifies the divergence of the network’s current performance from the ideal, via gradient descent [19]. To do so, the following steps are repeated until convergence: (1) forward propagation of information through the network; (2) evaluation of the loss function gradients with respect to the network parameters at the output layer; (3) backpropagation of these gradients to all previous layers; (4) parameter updates in the direction that maximally reduces the loss function. Forward propagation [step (1)] requires the aforementioned matrix multiplication, which maps information between layers, and a suitable nonlinear activation function, which is applied

individually to each neuron. While this nonlinearity has so far been mostly applied digitally in hybrid optical-electronic systems [9,20–22]—at the cost of repeatedly measuring and generating the optical state—recent work has also realized optical nonlinearities [18,23–25].

However, obtaining and backpropagating the loss-function gradients [(steps (2) and (3)] remains an outstanding problem in an optical setting. While backpropagating through the linear interconnection between layers is rather straightforward, as linear optical operations are naturally bidirectional, the nonlinearity of neurons is a challenge. This is because the backward-propagating signal must be modulated by the derivatives of the activation function of each neuron at its current input value, and these derivatives are not readily available in an ONN.

In 1987, Wagner *et al.* suggested that a feedforward ONN could be implemented and trained by using Fabry–Perot etalons to approximate the required forward and backward response of a sigmoid nonlinearity [26]. However, this backpropagation approach was never realized or even analyzed in detail, largely due to its inherent experimental complexity, with a subsequent ONN demonstration instead using digitally calculated errors [27]. A further approach to an optically trained feedforward network was proposed by Cruz-Cabrera *et al.* [28]. The authors used a highly nonstandard network architecture that transforms a “continuum of neurons” (a wavefront) as it passes through a nonlinear crystal using cross-phase modulation with a secondary “weight” beam. In a proof-of-concept experiment, the learning of two-bit logic was demonstrated.

An additional challenge is to map from the gradients with respect to the (platform-agnostic) weight matrices to the physical parameters that control these matrices in a specific ONN platform. In 2018, Hughes *et al.* [20] proposed a method to directly obtain the gradients of these control parameters by an additional forward-propagating step. However, this scheme assumes computing the derivatives of the activation functions digitally and applying them to the backpropagating signal electro-optically.

This work directly addresses the issue of optical backpropagation through nonlinear units in a manner that is both consistent with modern neural network architectures and compatible with leading ONN designs [9,20,29,30]. We consider an optical nonlinearity based on saturable absorption (SA) and show that, with the forward-propagating features and the backward-propagating errors taking the roles of pump and probe, respectively, backpropagation can be realized using only passive optical elements. Our method is effective and surprisingly simple, with the required optical operations for both forward and backward propagation realized using the same physical elements. Simulations with physically realistic parameters show that the proposed scheme can train networks to performance levels equivalent to state-of-the-art ANNs. When combined with optical calculation of the error term at the output layer via interference, this presents a path to the all-optical training of ONNs.

## 2. IMPLEMENTING OPTICAL BACKPROPAGATION

We begin by recapping the operation of a neural network before discussing optical implementations. Seeded with data at the

input layer ( $a^{(0)}$ ), forward-propagation maps the neuron activations from layer  $l - 1$  to the neuron inputs at layer  $l$  as

$$z_j^{(l)} = \sum_i w_{ji}^{(l)} a_i^{(l-1)}, \quad (1)$$

via a weight matrix  $w^{(l)}$ , before applying a nonlinear activation function individually to each neuron,  $a_j^{(l)} = g(z_j^{(l)})$  (with subscripts labelling individual neurons).

At the output layer, we evaluate the loss function,  $\mathcal{L}$ , and calculate its gradient with respect to the weights,

$$\frac{\partial \mathcal{L}}{\partial w_{ji}^{(l)}} = \frac{\partial \mathcal{L}}{\partial z_j^{(l)}} \frac{\partial z_j^{(l)}}{\partial w_{ji}^{(l)}} = \delta_j^{(l)} a_i^{(l-1)}, \quad (2)$$

where  $\delta_j^{(l)} \equiv \partial \mathcal{L} / \partial z_j^{(l)}$  is commonly referred to as the “error” at the  $j$ th neuron in the  $l$ th layer. From the chain rule, we have

$$\delta_j^{(l)} = \sum_k \frac{\partial \mathcal{L}}{\partial z_k^{(l+1)}} \frac{\partial z_k^{(l+1)}}{\partial z_j^{(l)}} = g'(z_j^{(l)}) \rho_j^{(l+1)}, \quad (3)$$

where  $\rho_j^{(l+1)} = \sum_k w_{kj}^{(l+1)} \delta_k^{(l+1)}$ . Given the error at the output layer, i.e.,  $\delta^{(L)}$ , which is calculated directly from the loss function, the errors  $\delta^{(L-1)}, \dots, \delta^{(1)}$  for all preceding layers are sequentially found using Eq. (3). These errors, as well as the activations  $a^{(l-1)}$  of all neurons, allow one to find the gradients [Eq. (2)] of the error function with respect to all the weights and hence apply the gradient descent.

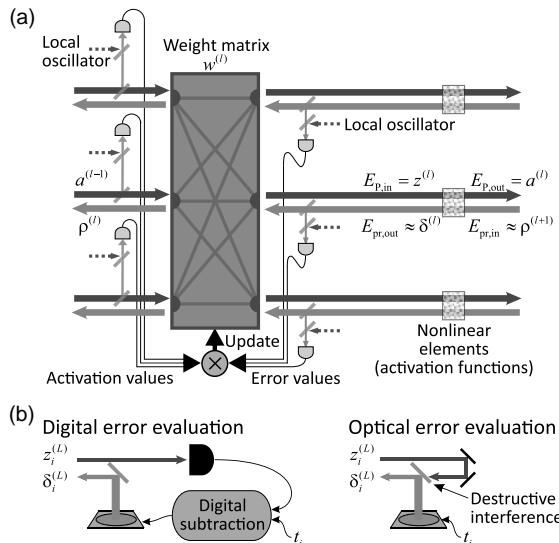
The transformation [Eq. (1)] is readily implemented as a linear optical (interferometric) operation, with the neurons represented by real-valued field amplitudes in different spatial modes [8]. An experimental realization of this operation, with the matrix and vector dimension up to 56, has recently been reported by our group [31]. Remarkably, calculating  $\rho^{(l+1)}$  in the right-hand side of the backpropagation in Eq. (3) involves the same weight matrix, meaning that it can be implemented by physical backward propagation of an optical signal through the same linear optical arrangement [27], as shown in Fig. 1(a). However, multiplying this signal by the derivative of the activation function,  $g'(z^{(l)})$ , is a challenge without invoking digital electronics.

To address this challenge, we require an optical implementation of the activation function with the following features: (i) nonlinear response for the forward input; (ii) linear response for the backward input; (iii) modulation of backward input with the derivative of the nonlinear function. While it is natural to use nonlinear optics for this purpose, it is difficult to satisfy the requirement that the unit must respond differently to forward- and backward-propagating light. Here, we show that this problem can be addressed using saturable absorption in the well-known pump-probe configuration.

Consider passing a strong pump,  $E_P$ , and a weak probe,  $E_{\text{pr}}$ , through a two-level medium (e.g., atomic vapor). The pump transmission is then a nonlinear function of the input,

$$E_{P,\text{out}} = g(E_{P,\text{in}}) = \exp \left( -\frac{\alpha_0/2}{1 + E_{P,\text{in}}^2} \right) E_{P,\text{in}}, \quad (4)$$

where  $\alpha_0$  is the resonant optical depth and all fields are assumed to be normalized by the saturation threshold. Figure 2(a) plots the pump transmission  $g(\cdot)$  at  $\alpha_0$  of 1 and 30. High optical depth induces strong nonlinearity in the unsaturated region,



**Fig. 1.** ONN with all-optical forward- and backward-propagation. (a) A single ONN layer that consists of weighted interconnections and an SA nonlinear activation function. The forward- (red) and backward-propagating (orange) optical signals, whose amplitudes are proportional to the neuron activations,  $a^{(l-1)}$ , and errors,  $\delta^{(l)}$ , respectively, are tapped off by beam splitters, measured by heterodyne detection and multiplied to determine the weight matrix update in Eq. (2). This multiplication can also be implemented optically, as discussed in the text. The final update of the weights, as well as the preparation of network input, is implemented electronically. (b) Error calculation at the output layer performed optically or digitally, as described in the text.

and a sufficiently strong pump renders the medium nearly transparent in the saturated region. A suitably weak probe, on the other hand, does not modify the transmissivity of the atomic media and hence experiences linear absorption with the absorption coefficient determined by the pump,

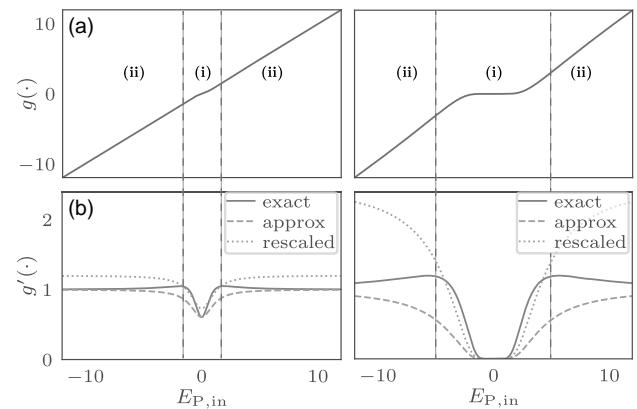
$$\frac{E_{p,out}}{E_{p,in}} = \exp\left(-\frac{\alpha_0/2}{1+E_{p,in}^2}\right). \quad (5)$$

Note that both beams are assumed to be resonant with the atomic transition and so, as the phase of the electric field is unchanged, we treat these as real-valued without loss of generality. Therefore, with the pump and probe taking the roles of the forward-propagating signal and backward-propagating error in an ONN, required features (i) and (ii) of our optical nonlinear unit are met.

Condition (iii), however, remains to be satisfied. The derivative of the pump transmission is

$$g'(E_{p,in}) = \left[1 + \frac{\alpha_0 E_{p,in}^2}{(1+E_{p,in}^2)^2}\right] \exp\left(-\frac{\alpha_0/2}{1+E_{p,in}^2}\right). \quad (6)$$

The derivatives at  $\alpha_0$  of 1 and 30 are plotted in Fig. 2(b). Our key insight is that, in many instances, the square-bracketed factor in Eq. (6) can be considered constant, in which case the backpropagation transmission of Eq. (5) is a good approximation of the desired response in Eq. (6) up to a constant factor. Feature (iii) is then satisfied because a constant scaling of the



**Fig. 2.** Saturable absorber response. (a) The transmission and (b) transmission derivative of an SA unit with optical depths of 1 (left) and 30 (right), as defined by Eqs. (4) and (6), respectively. Also shown in panel (b) are the actual probe transmissions given by Eq. (5), which approximate the derivatives, with and without the rescaling. The scaling factors are 1.2 (left) and 2.5 (right). In the amplitude region (i), the SA behaves as a linear absorber for weak input but then exhibits strong nonlinearity when the pump intensity approaches the saturation threshold. Region (ii) corresponds to strong saturation: the ground-state population is depleted, and the absorber is rendered transparent.

network gradients can be absorbed into the learning rate. This may appear as a coarse approximation, however; as we will see in the next section, it is only required to hold within the nonlinear region of the SA response, which is the case for our system [Fig. 2(b)].

The proposed scheme can be implemented on either integrated or free-space platforms. In the integrated setting, optical interference units that combine integrated phase-shifters and attenuators to realize intralayer weights have been demonstrated [9] as have, separately, on-chip SA through atomic vapor [32,33] and other nonlinear media [23,34]. A free-space implementation of the required matrix multiplication can be achieved using a spatial light modulator (SLM) [8,31] with the nonlinear unit provided by a standard atomic vapor cell. In the integrated case, an additional nontrivial step to map the weight gradients in Eq. (2) to suitable updates of the control parameters (i.e., phase-shifters and attenuator) is required; however, this challenge was recently addressed by Hughes *et al.* [20]. A free-space implementation, by contrast, has discrete blocks of SLM pixels directly controlling individual weights, so the update calculation is more straightforward.

Regardless of the chosen platform, passive optical elements can only implement weighted connections that satisfy conservation of energy. For networks with a single layer of nonlinear activations, this is not a practical limitation as the weight matrices can be effectively realized with normalized weights by correspondingly rescaling the neuron activations in the input layer. For deep networks with multiple layers, absorption through the vapor cell will reduce the field amplitude available to subsequent layers. This can be counteracted by interlayer amplification using, for example, semiconductor optical amplifiers [35].

In our proposed ONN, the only parts that require electronics are (a) real-valued homo- or heterodyne measurements of

the tapped-off neuron activations ( $a^{(l)}$ ) and error terms ( $\delta^{(l)}$ ) at each layer, (b) generating the network input and reference beams, and (c) updating the weights. In practice, the update (c) is calculated not for each individual training set element but as the average over multiple elements (a “mini-batch” or a training epoch); hence, the speed of this operation is not critical for the ONN performance. Generating the inputs, (b), and targets is decoupled from the calculation performed by the ONN and requires fast optical modulators, which are abundant on the market.

Finally, the measurements, (a), must be followed by calculating the product  $\delta_j^{(l)} a_i^{(l-1)}$  and averaging over the minibatch. This operation can be implemented using electronic gate arrays. For a network with  $L$  layers of  $N$  neurons, this requires  $2LN$  measurements and  $LN^2$  offline multiplications. Alternatively, the multiplication can be realized by direct optical interference of the two signals with each other, followed by intensity measurement. The optical multiplication would require phase stability of the setup and the additional overhead of  $2LN^2$  photodetectors but eliminate the need for reference beams and offline multiplications. Additionally, calculating the new weight matrices from the data acquired through these detectors will require  $\sim LN^2$  operations, which would need to be performed once per epoch. Note that these operations may introduce a performance bottleneck due to the limited refresh rate of the modulators.

Although the activation and error signals are frequency-degenerate in our scheme, their counterpropagating configuration makes them easily distinguishable at the detection stage. Furthermore, the nonlinear unit operation is not affected by the relative phase of the two counterpropagating signals.

The primary latencies associated with the optical propagation of the signal in the ONN are due to the bandwidths of the SAs and intralayer amplifiers. Further processing speed limitations are present in the photodetection and multiplication of  $\delta_j^{(l)} a_i^{(l-1)}$  as well as conversion of the computed weight matrix gradients to their actuators within the ONN [20]. This latter conversion, however, occurs once per training batch, so this limitation can be amortized by using large batches.

The remaining, not yet discussed, element of the ONN training is the calculation and reinjection of the error  $\delta^{(L)}$  at the output layer, to initiate backpropagation. To implement this optically, we train the ONN with the mean-squared-error loss function,

$$\mathcal{L} = \sum_i \frac{1}{2} (z_i^{(L)} - t_i)^2, \quad (7)$$

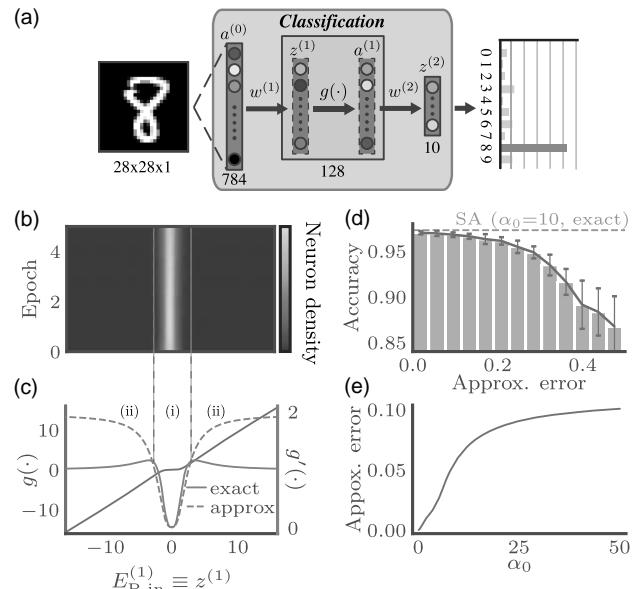
where  $t_i$  is the target value for the  $i$ th output neuron. This loss function implies  $\delta_i^{(L)} = \partial \mathcal{L} / \partial z_i^{(L)} = z_i^{(L)} - t_i$ , which is calculable by interference of the network outputs with the target outputs on a balanced beam-splitter. This approach to error calculation is illustrated in the right panel of Fig. 1(b), whereas the left panel shows the standard approach in which the errors are calculated offline (electronically).

### 3. EXAMINING APPROXIMATION ERRORS

To investigate our proposed backpropagation scheme and, in particular, how our approximated derivatives affect network

performance, we consider the canonical ML task of image classification. Our first set of numerical experiments classifies images of handwritten digits from 0 to 9. We use the MNIST [36] data set, which contains grey-scale bitmaps of size  $28 \times 28$  fed into the input layer of the ONN. The output layer contains 10 neurons whose target values are 0 or 1 dependent on the digit encoded in the bitmap (“one-hot encoding”). In this section, we use a network architecture with a single 128-neuron hidden layer, as shown in Fig. 3(a). Further details of the networks, training, and calculation of the accuracy metric for all experiments presented in this work can be found in Appendix A.

Initially, we consider the activation function to be provided by SA with an optical depth of  $\alpha_0 = 10$ . For the chosen network architecture, this provides  $97.3\% \pm 0.1\%$  classification accuracy after training, with no difference in performance regardless of whether the true derivatives in Eq. (6) or the optically obtainable derivative approximations in Eq. (5) are used. From Fig. 3(b), we can see that, during training, the neurons are primarily distributed in the unsaturated region of the SA activation function. This is a consequence of the fact that the expressive capacity of neural networks arises from the nonlinearity of its neurons. Therefore, to train the network, the optically obtained derivatives need to approximate the exact derivatives (up to a fixed scaling as previously discussed) in only this nonlinear region.



**Fig. 3.** Effects of imperfect approximation of the activation function derivative. (a) Feed-forward neural network architecture using a single hidden layer of 128 neurons. (b) Distribution of neuron inputs ( $E_{P,in}^{(1)} \equiv z^{(1)}$ ), which is concentrated in the unsaturated region (1) of the SA activation function,  $g(\cdot)$ . As a result, the approximation error in the linear region (2) is less impactful on the training. (c) The transmission of an SA unit with  $\alpha_0 = 10$ , along with the exact and (rescaled for easier comparison) optically approximated transmission derivatives. (d) Performance loss associated with approximating activation function derivatives  $g'(\cdot)$  with random functions, plotted as a function of the approximation error, for  $\alpha_0 = 10$  (see Appendix B for details). (e) Average error of the derivative approximation in Eq. (5) as a function of the optical depth of an SA nonlinearity.

It is interesting to investigate how training is affected by imprecision in the derivatives used. To this end, we evaluate the network performance by replacing the derivative  $g'(\cdot)$  with random functions of varying similarity to the true derivative within the nonlinear region (the quantitative measure,  $S$ , of the similarity is defined in Appendix B). From Fig. 3(d), we see that the performance appears robust to approximation errors, defined as  $1-S$ , of up to  $\sim 15\%$ . We explain this potentially surprising observation by noting that gradient descent will converge even if the update vector deviates from the direction toward the exact minimum of the loss function, so long as this deviation is not too significant.

In the case of SA, i.e., when the approximate derivatives given by Eq. (5) are used, this error saturates at  $\sim 10\%$  for increasing optical depth [see Fig. 3(e)] so no significant detrimental effect on the training accuracy can be expected. These results suggest that our scheme would still be effective in a noisy experimental setting, as further discussed in Appendix D, and that the approach studied here may function well for a broad range of optical nonlinearities.

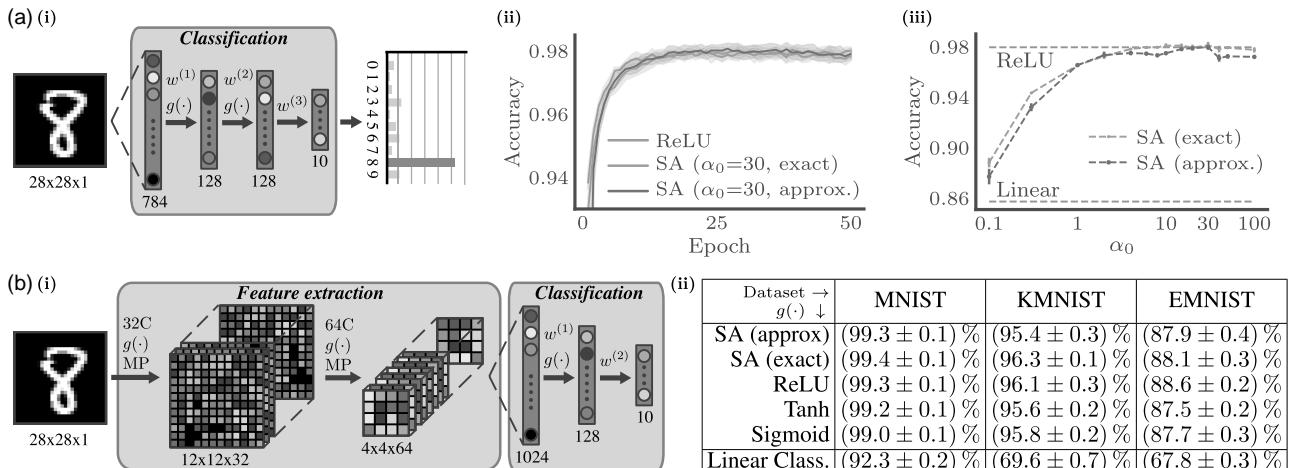
#### 4. CASE STUDY: IMAGE CLASSIFICATION

Thus far, we have only used a simple network architecture to examine our derivative approximation; however, we now consider how ONNs with SA nonlinearities are compared with state-of-the-art ANNs. To do this, we use deeper network architectures for a range of image classification tasks. To obtain a comparison benchmark, we computationally train ANNs with equivalent architectures using standard best practices. Concretely, for ANNs we use ReLU (rectified linear unit) activation functions, defined as  $g_{\text{ReLU}}(z) = \max(0, z)$ , and the categorical cross-entropy loss function, which is defined as

$\mathcal{L} = -\sum_i t_i \log(p_i)$ , where  $p_i = \exp(z_i^{(L)}) / \sum_k \exp(z_k^{(L)})$  is the softmax probability distribution of the network output (see Appendix A for a discussion of the different choices of loss function for ANNs and ONNs).

To begin, we use a network with two 128-neuron hidden layers, as shown in Fig. 4(a)(i), and once again consider the MNIST data set. Figure 4(a)(ii) compares the simulated performance of the optical and benchmark networks. The ReLU-based classifier achieves an accuracy of  $98.0\% \pm 0.2\%$ , which provides an approximate upper bound on the achievable performance of this network architecture for the chosen task [37]. An optical network with an optical depth of  $\alpha_0 = 30$  [trained with approximate activation function derivatives in Eq. (5) obtained in optical backpropagation through the SA] exactly matches this level of performance with  $98.0\% \pm 0.2\%$  classification accuracy. As an additional benchmark, we train the optical network using the exact derivative in Eq. (6) of the activation function, obtaining a similar accuracy of  $98.1\% \pm 0.3\%$ . The convergence speed to near-optimum performance during training is unchanged across all of these networks.

Figure 4(a)(iii) shows the trained performance of optical networks as a function of the optical depth, which essentially determines the degree of nonlinearity of the transmission function. As  $\alpha_0 \rightarrow 0$ , our network can only learn linear functions of the input, which restricts the classification accuracy to  $85.7\% \pm 0.4\%$ . For larger optical depths, the performance of the network improves, with the strong performance observed at  $\alpha_0 = 1$  increasing to near optimal levels once  $\alpha_0 \geq 10$ , which is readily obtainable experimentally. Eventually, for  $\alpha_0 \geq 30$ , we start to see the performance of the approximated derivatives reduced, although high accuracy is still obtained. This can be attributed to the increasing approximation errors associated with high optical depths [see Fig. 3(e)], which, as



**Fig. 4.** Performance on image classification. (a) (i) The fully connected network architecture. (ii) Learning curves for the SA [with either exact derivatives in Eq. (6) of the activation function or their approximation in Eq. (5)] and benchmark ReLU networks. (iii) The final classification accuracy achieved as a function of the optical depth,  $\alpha_0$ , of the SA cell. (b) (i) The convolutional network architecture. Sequential convolution layers of 32 and 64 channels convert a  $28 \times 28$  pixel image into a 1024-dimensional feature vector, which is then classified (into  $N_C = 10$  classes for MNIST and KMNIST, and  $N_C = 47$  classes for EMNIST) by fully connected layers. Pooling layers are not shown for simplicity. (ii) Classification accuracy of convolutional networks when using various activation functions. The same deep network architecture is applied to all data sets, but the SA networks use mean-pooling, while the benchmark networks use max-pooling. The last row shows the performance of a simple linear classifier as a baseline.

previously discussed, accumulate in the deeper network architecture. In free-space implementations with saturated atomic vapor cells, the optical depth can be dynamically tuned via cell temperature. For semiconductor absorbers in integrated settings, the optical depth is related to the material thickness and/or the density of dopants.

To probe the limits of the achievable performance using SA nonlinearities and optical backpropagation, we also consider the more challenging Kuzushiji-MNIST [38] (KMNIST) and extended-MNIST [39] (EMNIST) data sets. For these applications, we use a deep network architecture with convolutional layers (see Appendix A for details), as illustrated in Fig. 4(b)(i), which significantly increases the achievable classification accuracy to a level approaching the state-of-the-art. While not the focus of this work, we emphasize that convolutional operations are readily achievable with optics. Current research into convolutional ONNs either directly leverages imaging systems [40] or decomposes the required convolution into optical matrix multiplication [41–43].

In addition to convolutional layers, convolutional neural networks also contain pooling layers, which locally aggregate neuron activations. The common implementation of these is max-pooling; however, this operation does not readily translate to an optical setting. Therefore, for ONNs, we deploy mean-pooling, where the activation of neurons is locally averaged, which is a straightforward linear optical operation. In contrast, our benchmark ANNs utilize max-pooling.

Figure 4(b)(ii) compares the obtained performance with SA nonlinearities (with  $\alpha_0 = 10$ ) to that achieved with benchmark ANNs that use various standard activation functions. We see an equivalent level of performance, despite the approximation in the backpropagation phase. This result suggests that all-optical backpropagation can be utilized to train sophisticated networks to state-of-the-art levels of performance.

## 5. DISCUSSION

This work presents an effective and surprisingly simple approach to achieving optical backpropagation through nonlinear units in a neural network, an outstanding challenge in the pursuit of truly all-optical networks. With our scheme, the information propagates through the network in both directions without interconversion between optical and electronic form. The role of digital electronics is reduced to the preparation of the network input, photodetection, and updating the network parameters. In these elements of the network, the conversion speed is not critical, particularly for large batches of training data. A detailed estimate of the energy efficiency and computation speed of the optically trained neural network is presented in Appendix C.

As compared with offline training, optical training is more robust against experimental imperfections, since such imperfections are automatically included and counteracted during the training process. As an illustration, numerical simulation results of optical training in noisy experimental settings are presented in Appendix D.

The scheme is compatible with a variety of ONN platforms. We also anticipate that a broader class of nonlinear optical phenomena can be used to implement the activation function.

For example, one could consider directly using saturation of intralayer amplifiers for this role, circumventing the need for SA units entirely. A preliminary numerical experiment to this effect is discussed in Appendix E. Our scheme may also be applicable to the optical Kerr nonlinearity as proposed in Ref. [30] for training diffractive neural networks, albeit with the added complexity of operating with complex-valued field amplitudes and weights. Therefore, as well as presenting a path toward the end-to-end optical training of neural networks, this work sets out an important consideration for nonlinearities in the design of analog neural networks of any nature.

## APPENDIX A. NETWORK DETAILS

### 1. Image Data Sets

We consider three different data sets, all containing  $28 \times 28$  pixel grey-scale images: MNIST [36], Kuzushiji-MNIST (KMNIST) [38], and extended-MNIST (EMNIST) [39]. MNIST corresponds to handwritten digits from 0 to 9; KMNIST contains 10 classes of handwritten Japanese cursive characters; and we use the EMNIST balanced data set, which contains 47 classes of handwritten digits and letters. MNIST and KMNIST have 70,000 images in total, split into 60,000 training and 10,000 test instances. EMNIST has 131,600 images, with 112,800 (18,800) training (test) instances. For all data sets, the training and testing sets have all classes equally represented.

### 2. Network Architectures

The fully connected network we train to classify MNIST [corresponding to the results in Fig. 4(a)] first unrolls each image into a 784-dimensional input vector, before two 128-neuron hidden layers and a 10-neuron output layer.

The convolutional network depicted in Fig. 4(b)(i) has two convolutional layers of 32 and 64 channels, respectively. Each layer convolves the input with  $5 \times 5$  filters (with a stride of 1 and no padding), followed by a nonlinear activation function and finally a pooling operation (with both kernel size and stride of 2). After the convolutional network, classification is carried out by a fully connected network with a single 128-neuron hidden layer and  $N_C$  neuron output layer, where  $N_C$  is the number of classes in the target data set.

Multilayer ONNs are assumed to have the same optical depth of their saturable absorbers in all layers.

### 3. Network Loss Function

As stated in the main text, we train ONNs using the mean-squared-error (MSE) loss function, whereas the ANN baselines use categorical cross-entropy (CCE). This choice was made as the gradients of MSE loss are readily calculable in an optical setting, whereas the softmax operation in CCE would require offline calculation. However, our ANNs use CCE, as this is the standard choice for classification problems in the deep learning community. For completeness, we retrained our ANN baselines for MNIST classification using MSE. The fully connected classifier [Fig. 4(a)(i)] provided a classification accuracy of  $98.0\% \pm 0.2\%$ , while the convolutional classifier [Fig. 4(a)(ii)], using ReLU nonlinearities, scored  $99.5\% \pm 0.1\%$ . In both cases, the performance of MSE is essentially equivalent to that of CCE.

#### 4. Network Training

All networks are trained with a minibatch size of 64. We used the Adam optimizer with a learning rate of  $5 \times 10^{-4}$ , independent of the optical depth of the SA. For each network, the test images of the target data set are split evenly into a “validation” set and a “test” set. After every epoch, the performance of the network is evaluated on the held-out “validation” images. The best ONN parameters found over training are then used to verify the performance on the “test” set. Therefore, learning curves showing the performance during training [i.e., Fig. 4(a)(ii)] are plotted with respect to the “validation” set, with all other reported results corresponding to the “test” set. The fully connected networks were trained on MNIST for 50 epochs. The convolutional networks are trained for 20 epochs when using ReLU, Tanh, or Sigmoid nonlinearities and 40 epochs when using SA nonlinearities.

Training performance is empirically observed to be sensitive to the initialization of the weights, which we ascribe to the small derivatives away from the nonlinear region of the SA response curve. For low optical depths,  $\alpha_0 \leq 30$ , all layers are initialized as a normal distribution of width 0.1 centered around 0. For higher optical depths, the weights of the fully connected ONN shown in Fig. 4(a) are initialized to a double-peaked distribution comprised of two normal distributions of width 0.15 centered at  $\pm 0.15$ . We do not constrain our weight matrices during training because, as discussed in the main text, conservation of energy can always be satisfied by rescaling the input power or output threshold for the first and last linear transformation and using intralayer amplifiers in deeper architectures.

For all images, the input is rescaled to be between 0 and 1 (which practically would correspond to  $0 \leq E_{p,\text{in}}^{(0)} \leq 1$ ) when passed to a network with computational nonlinearities (i.e., ReLU, Sigmoid, or Tanh). Due to “absorption” in networks with SA nonlinearities, we empirically observe that rescaling the input data to higher values results in faster convergence when training convolutional networks with multiple hidden layers. Therefore, the fully connected networks in Fig. 4(a) use inputs between 0 and 1, and the convolutional networks in Fig. 4(b) use inputs normalized between 0 and 5 (15) for  $\alpha_0 \leq 10$  ( $\alpha_0 > 10$ ).

#### APPENDIX B. CALCULATION OF THE DERIVATIVE APPROXIMATION ERROR

As discussed in the main text, we approximate the true derivatives  $g'(\cdot)$  of the activation functions by random functions  $f(\cdot)$  to test the effect of the approximation error on training. Here, we discuss how these functions are generated and how the similarity measure is defined.

The response of a saturable absorption nonlinearity can be considered in two regimes, i.e., nonlinear (unsaturated) and linear (saturated), which are labeled (i) and (ii) in Fig. 2, respectively. During the network training, the neuron input values ( $z_j^{(l)}$ ) are primarily distributed in the nonlinear region, as seen in Fig. 3(b) and discussed in the main text. Therefore, we model the neuron input as a Gaussian distribution within this region:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{z^2}{2\sigma^2}\right), \quad (\text{B1})$$

where  $2\sigma$  is the width of region (1). We then define the similarity as the reweighted normalized scalar product between the accurate and approximate derivatives,

$$S = \frac{|\int f(z)g'(z)p(z)dz|^2}{\int [f(z)]^2p(z)dz \cdot \int [g'(z)]^2p(z)dz}. \quad (\text{B2})$$

According to the Cauchy–Schwarz inequality,  $S$  is bounded by 1 and therefore so is the average approximation error,  $1-S$ .

To obtain the results in Figs. 3(d) and 3(e), we generate 200 random functions for  $f$ , with different approximation errors. We first generate an array of pseudo-random numbers ranging from 0 to 1, concatenate it with the flipped array to make them symmetric like the derivative  $g'(\cdot)$ , and then use shape-preserving interpolation to obtain a smooth and symmetric random function. The network is then trained once with each of the generated  $f$ .

#### APPENDIX C. OPTICAL POWER CONSUMPTION AND COMPUTATION SPEED

The optical power consumption in an ONN depends on the network architecture and implementation details. For concreteness, we consider a fully connected network with  $N = 1000$  units per layer, with SA optical nonlinearities implemented on the  $^{87}\text{Rb}$  D<sub>2</sub> line. Recalling Fig. 3(b) from the main text, we note that, during training, the input power to each neuron is typically restricted to the unsaturated region, (i), of the nonlinearity response. For the SA nonlinearities we consider, the saturation intensity is given by [44]

$$I_{\text{sat}} = \frac{\hbar\omega\Gamma}{2\sigma_0} = 16.6 \mu\text{W} \cdot \text{mm}^{-2}, \quad (\text{C1})$$

where  $\Gamma = 2\pi \times 6 \text{ MHz}$  is the natural linewidth, and  $\sigma_0 = 3\lambda^2/(2\pi)$  is the resonant absorption cross section. For beams with a waist of  $w_0 = 100 \mu\text{m}$ , this corresponds to a saturation power of  $P_{\text{sat}} \approx 500 \text{ nW}$  per neuron and total SA input power for 1000 units on the order of 500  $\mu\text{W}$ .

To saturate the SA, the optical pulse needs to be longer than the excited state lifetime  $\Gamma^{-1} = 26 \text{ ns}$ . The energy cost of a single forward pass through the network is then  $\sim 10 \text{ pJ}$ , and the backpropagation energy cost is negligible. Since a single interlayer transition involves a VMM with  $N^2$  multiplications, one can estimate the energy cost per multiply-accumulate operation to be  $\sim 10 \text{ aJ}$ . In an integrated setting, the saturation powers are higher, but the pulse durations are proportionally shorter, so similar energy costs can be expected. This is, of course, an idealized estimate, which does not include peripheral energy costs in powering and sustaining the instruments and stabilizing the system. Hence, the actual power consumption can be expected to be at least an order of magnitude higher. For comparison, today’s electronic processors like CPUs and GPUs have energy costs on a scale of 0.1 to 1  $\text{nJ}$  per operation.

The computation speed of the ONN is determined by the response time of the SA units and amplifiers as well as the speed of optical modulators. The response time of atomic-based SA is tens of nanoseconds and that of semiconductor SA is on the order of picoseconds. Response time of optical amplifiers is

on a similar time scale. The speed of the optical modulator preparing the network input is more likely to be the bottleneck in the near future. The bandwidths of SLM, thermo-optic modulator, and electro-optic modulator are on the order of 10 kHz [31], 100 kHz [9], and 10 GHz [45], respectively. Adoption of an ultrafast electro-optic modulator in an ONN with 1000 neurons per layer would perform  $10^{16}$  operations per second. Assuming the energy consumption of 1 mW as estimated above, the energy efficiency would be  $10^{19}$  operations per second per watt. For comparison, a higher-end consumer GPU has a computation speed of  $\sim 10^{12}$  operations per second, with the energy efficiency of  $\sim 10^{10}$  operations per second per watt, and neuromorphic electronics has an energy efficiency below  $10^{14}$  operations per second per watt [13].

#### APPENDIX D. OPTICAL TRAINING WITH EXPERIMENTAL IMPERFECTIONS

A practical ONN will exhibit noise and errors arising from background scattering, nonideal optical multiplication or interference, and digitization error of electronic signals. Here, we investigate the robustness of our scheme against these imperfections. We adopt the network structure of Fig. 3 and perform simulation on the MNIST data set.

In the first series of tests, we add a certain amount of random Gaussian noise to the activation outputs  $a_i^{(l)}$  and error fields  $\delta_i^{(l)}$ . We define the noise level as the ratio between the standard deviation of noise and signal:  $\sigma_{\text{noise}}/\sigma_{\text{signal}}$ . The top three rows in Table 1 show the training result. The classification accuracy decreases mildly from 97.3% to 95.8% with 10% noise.

Our second series of tests consists in randomly scaling the activation output and error beams: each mode has a fixed scaling factor, and all the scaling factors are sampled from a Gaussian distribution  $N(1, \sigma)$ . This models the nonuniform losses of different spatial modes. From Table 1 (fourth to sixth rows), we see that such imperfection causes no performance degradation because the weights are automatically rescaled during training to counteract such deviation.

We further consider the digitization error of weights, activation, and error beams, since they are usually electronically controlled or read out with limited-bitwidth analog-to-digital or digital-to-analog converters. The seventh to ninth rows of Table 1 show that the training is sensitive to the bitwidth limitations, and the accuracy drops to 96.0% with 8-bit precision and 93.5% with 6-bit precision. Therefore, in an actual system, one should use at least 8-bit controls to preserve high accuracy

**Table 1. ONN Training with Experimental Imperfections**

Noise Level	Deviation Level	Bitwidth	Accuracy
0%	0%	32	97.3% $\pm$ 0.1%
5%	0%	32	96.7% $\pm$ 0.2%
10%	0%	32	95.8% $\pm$ 0.3%
0%	5%	32	97.1% $\pm$ 0.3%
0%	10%	32	97.1% $\pm$ 0.1%
0%	20%	32	97.1% $\pm$ 0.3%
0%	0%	8	96.0% $\pm$ 0.4%
0%	0%	6	93.5% $\pm$ 0.8%
5%	10%	8	96.1% $\pm$ 0.2%

of the network, which can be readily achieved. Note that high-performance ONNs with bitwidths as low as 2 to 4 bits have been proposed [46].

Finally, the last row of Table 1 shows that classification accuracy as high as 96.1% can still be achieved for a practical 8-bit system with the combined effect of 5% random noise and up to 10% deviation.

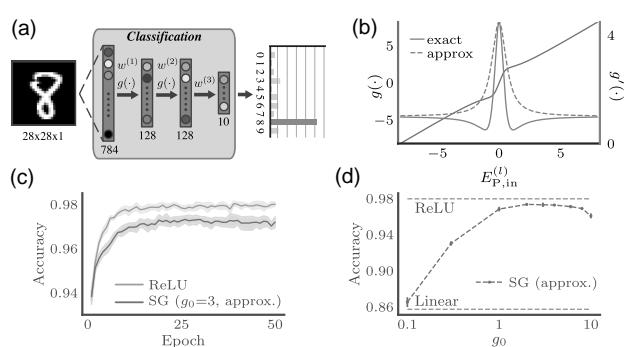
#### APPENDIX E. OPTICAL BACKPROPAGATION WITH SATURABLE GAIN

In optical amplifiers, saturable gain (SG) takes place when a sufficiently high input power depletes the excited state of the gain medium. In a two-level system, this process can be described similarly to saturable absorption by simply replacing the optical depth term  $-a_0$  in Eq. (4) with a positive gain factor  $g_0$ . The transmission in the forward direction is plotted and compared with both the theoretically exact and optically obtained transmission derivatives in Fig. 5(b) with  $g_0 = 3$ . The derivative curves have the inverted shapes of the SA derivative curves.

The optically obtained derivatives also appear to be a reasonable approximation of the exact gradient. To examine this, we replace the SA nonlinearity with SG nonlinearity in the fully connected network, as shown in Fig. 4(a), and repeat the optical training simulation. The MNIST image classification performance is shown in Figs. 5(c) and 5(d). High accuracy can be achieved with a gain factor as small as 1, and the best result scores  $97.3\% \pm 0.1\%$  at  $g_0 = 3$ , slightly lower than that of the benchmark ReLU network and SA-based ONN. Since the derivative approximation error of the SG nonlinearity is the same as that of the SA nonlinearity, the performance degradation is mainly attributed to the nonlinearity itself; however, higher performance may be achievable through careful hyperparameter tuning.

#### APPENDIX F. CODE AVAILABILITY

Source code is available in Code 1, Ref. [47].



**Fig. 5.** Optical backpropagation through saturable gain (SG) nonlinearity. (a) Fully connected network architecture, which is the same as Fig. 4(a) except for the nonlinearity. (b) Transmission and transmission derivatives of the SG unit with gain factor  $g_0 = 3$ . (c) Learning curves for the SG-based ONN and benchmark ReLU networks. (d) The final classification accuracy achieved as a function of the gain.

**Funding.** Russian Science Foundation (19-71-10092).

**Acknowledgment.** X. G. acknowledges funding from the University of Electronic Science and Technology of China. A. L. thanks W. Andregg for introducing him to ONNs.

**Disclosures.** The authors declare no conflicts of interest.

<sup>a</sup>These authors contributed equally to this work.

## REFERENCES

- E. Cambria and B. White, "Jumping NLP curves: a review of natural language processing research," *IEEE Comput. Intell. Mag.* **9**, 48–57 (2014).
- W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: a comprehensive review," *Neural Comput.* **29**, 2352–2449 (2017).
- D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. Van Den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, and S. Dieleman, "Mastering the game of go with deep neural networks and tree search," *Nature* **529**, 484–489 (2016).
- J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proceedings of the 34th International Conference on Machine Learning* (2017), Vol. **70**, pp. 1263–1272.
- G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, "Neural-network quantum state tomography," *Nat. Phys.* **14**, 447–450 (2018).
- K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.* **2**, 359–366 (1989).
- G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control Signals Syst.* **2**, 303–314 (1989).
- P. N. Tamura and J. C. Wyant, "Two-dimensional matrix multiplication using coherent optical techniques," *Opt. Eng.* **18**, 182198 (1979).
- Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
- J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–214 (2019).
- L. De Marinis, M. Cococcioni, P. Castoldi, and N. Andriolli, "Photonic neural networks: a survey," *IEEE Access* **7**, 175827 (2019).
- A. N. Tait, T. F. De Lima, E. Zhou, A. X. Wu, M. A. Nahmias, B. J. Shastri, and P. R. Prucnal, "Neuromorphic photonic networks using silicon photonic weight banks," *Sci. Rep.* **7**, 7430 (2017).
- B. J. Shastri, A. N. Tait, T. Ferreira de Lima, M. A. Nahmias, H. T. Peng, and P. R. Prucnal, "Principles of neuromorphic photonics," in *Encyclopedia of Complexity and Systems Science*, R. A. Meyers, ed. (Springer, 2018).
- Y. Abu-Mostafa and D. Psaltis, "Optical neural computers," *Sci. Am.* **256**, 88–95 (1987).
- S. Jutamulia and F. Yu, "Overview of hybrid optical neural networks," *Opt. Laser Technol.* **28**, 59–72 (1996).
- J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, "Reinforcement learning in a large-scale photonic recurrent neural network," *Optica* **5**, 756–760 (2018).
- X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
- Y. Zuo, B. Li, Y. Zhao, Y. Jiang, Y.-C. Chen, P. Chen, G.-B. Jo, J. Liu, and S. Du, "All-optical neural network with nonlinear activation functions," *Optica* **6**, 1132–1137 (2019).
- Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
- T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, "Training of photonic neural networks through *in situ* backpropagation and gradient measurement," *Optica* **5**, 864–871 (2018).
- I. A. D. Williamson, T. W. Hughes, M. Minkov, B. Bartlett, S. Pai, and S. Fan, "Reprogrammable electro-optic nonlinear activation functions for optical neural networks," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700412 (2020).
- M. M. P. Fard, I. A. Williamson, M. Edwards, K. Liu, S. Pai, B. Bartlett, M. Minkov, T. W. Hughes, S. Fan, and T. A. Nguyen, "Experimental realization of arbitrary activation functions for optical neural networks," *Opt. Express* **28**, 12138–12148 (2020).
- Z. Cheng, H. K. Tsang, X. Wang, K. Xu, and J.-B. Xu, "In-plane optical absorption and free carrier absorption in graphene-on-silicon waveguides," *IEEE J. Sel. Top. Quantum Electron.* **20**, 4400106 (2013).
- A. Jha, C. Huang, and P. R. Prucnal, "Reconfigurable all-optical nonlinear activation functions for neuromorphic photonics," *Opt. Lett.* **45**, 4819–4822 (2020).
- M. Miscuglio, A. Mehrabian, Z. Hu, S. I. Azzam, J. George, A. V. Kildishev, M. Pelton, and V. J. Sorger, "All-optical nonlinear activation function for photonic neural networks," *Opt. Mater. Express* **8**, 3851–3863 (2018).
- K. Wagner and D. Psaltis, "Multilayer optical learning networks," *Appl. Opt.* **26**, 5061–5076 (1987).
- D. Psaltis, D. Brady, and K. Wagner, "Adaptive optical networks using photorefractive crystals," *Appl. Opt.* **27**, 1752–1759 (1988).
- A. A. Cruz-Cabrera, M. Yang, G. Cui, E. C. Behrman, J. E. Steck, and S. R. Skinner, "Reinforcement and backpropagation training for an optical neural network using self-lensing effects," *IEEE Trans. Neural Netw.* **11**, 1450–1457 (2000).
- T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, and Q. Dai, "Large-scale neuromorphic optoelectronic computing with a reconfigurable diffractive processing unit," *arXiv:2008.11659* (2020).
- T. Zhou, L. Fang, T. Yan, J. Wu, Y. Li, J. Fan, H. Wu, X. Lin, and Q. Dai, "*In situ* optical backpropagation training of diffractive optical neural networks," *Photon. Res.* **8**, 940–953 (2020).
- J. Spall, X. Guo, T. D. Barrett, and A. I. Lvovsky, "Fully reconfigurable coherent optical vector-matrix multiplication," *Opt. Lett.* **45**, 5752–5755 (2020).
- W. Yang, D. B. Conkey, B. Wu, D. Yin, A. R. Hawkins, and H. Schmidt, "Atomic spectroscopy on a chip," *Nat. Photonics* **1**, 331–335 (2007).
- R. Ritter, N. Gruhler, W. Pernice, H. Kübler, T. Pfau, and R. Löw, "Atomic vapor spectroscopy in integrated photonic structures," *Appl. Phys. Lett.* **107**, 041101 (2015).
- Q. Bao, H. Zhang, Y. Wang, Z. Ni, Y. Yan, Z. X. Shen, K. P. Loh, and D. Y. Tang, "Atomic-layer graphene as a saturable absorber for ultrafast pulsed lasers," *Adv. Funct. Mater.* **19**, 3077–3083 (2009).
- M. J. Connelly, *Semiconductor Optical Amplifiers* (Springer, 2007).
- Y. LeCun, C. Cortes, and C. Burges, and ATT Labs, "MNIST handwritten digit database," <http://yann.lecun.com/exdb/mnist> (2010).
- Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE* **86**, 2278–2324 (1998).
- T. Clanuwat, M. Bober-Irizar, A. Kitamoto, A. Lamb, K. Yamamoto, and D. Ha, "Deep learning for classical Japanese literature," *arXiv:1812.01718* (2018).
- G. Cohen, S. Afshar, J. Tapson, and A. van Schaik, "EMNIST: an extension of MNIST to handwritten letters," *arXiv:1702.05373* (2017).
- J. Chang, V. Sitzmann, X. Dun, W. Heidrich, and G. Wetzstein, "Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification," *Sci. Rep.* **8**, 12324 (2018).
- H. Bagherian, S. Skirlo, Y. Shen, H. Meng, V. Ceperic, and M. Soljačić, "On-chip optical convolutional neural networks," *arXiv:1808.03303* (2018).
- S. Xu, J. Wang, R. Wang, J. Chen, and W. Zou, "High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays," *Opt. Express* **27**, 19778–19787 (2019).
- R. Hamerly, L. Bernstein, A. Sludds, M. Soljačić, and D. Englund, "Large-scale optical neural networks based on photoelectric multiplication," *Phys. Rev. X* **9**, 021032 (2019).
- D. A. Steck, "Rubidium 87 D line data," <http://steck.us/alkalidata/rubidium87numbers.pdf>.

45. C. Wang, M. Zhang, X. Chen, M. Bertrand, A. Shams-Ansari, S. Chandrasekhar, P. Winzer, and M. Lonoar, "Integrated lithium niobate electro-optic modulators operating at CMOS-compatible voltages," *Nature* **562**, 101–104 (2018).
46. M. Miscuglio and V. J. Sorger, "Photonic tensor cores for machine learning," *Appl. Phys. Rev.* **7**, 031404 (2020).
47. T. D. Barrett, "Tomdbar/all-optical-neural-networks: publication code (Version v1.0.0)," <https://doi.org/10.5281/zenodo.4059672> (2020).

# PHOTONICS Research

## Learning to recognize misaligned hyperfine orbital angular momentum modes

XIAO WANG,<sup>1</sup> YUFENG QIAN,<sup>1</sup> JINGJING ZHANG,<sup>1</sup> GUANGDONG MA,<sup>1</sup> SHUPENG ZHAO,<sup>1</sup> RUIFENG LIU,<sup>1,\*</sup> HONGRONG LI,<sup>1</sup> PEI ZHANG,<sup>1</sup> HONG GAO,<sup>1</sup> FENG HUANG,<sup>2,3</sup> AND FULI LI<sup>1</sup>

<sup>1</sup>Shaanxi Province Key Laboratory of Quantum Information and Quantum Optoelectronic Devices, School of Physics, Xi'an Jiaotong University, Xi'an 710049, China

<sup>2</sup>School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China

<sup>3</sup>e-mail: huangf@fzu.edu.cn

\*Corresponding author: ruifeng.liu@mail.xjtu.edu.cn

Received 20 October 2020; revised 4 January 2021; accepted 26 January 2021; posted 29 January 2021 (Doc. ID 412965); published 15 March 2021

Orbital angular momentum (OAM)-carrying beams have received extensive attention due to their high-dimensional characteristics in the context of free-space optical communication. However, accurate OAM mode recognition still suffers from reference misalignment of lateral displacement, beam waist size, and initial phase. Here we propose a deep-learning method to exquisitely recognize OAM modes under misalignment by using an alignment-free fractal multipoint interferometer. Our experiments achieve 98.35% recognizing accuracy when strong misalignment is added to hyperfine OAM modes whose Bures distance is 0.01. The maximum lateral displacement we added with respect to the perfectly on-axis beam is about  $\pm 0.5$  beam waist size. This work offers a superstable proposal for OAM mode recognition in the application of free-space optical communication and allows an increase of the communication capacity. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.412965>

### 1. INTRODUCTION

It is well known that the orbital angular momentum (OAM) of photons was discovered by Allen *et al.* [1] in 1992. Since the topological charge  $\ell$  can be any integer, the OAM-carrying beams have countless orthogonal eigenstates, which allows them to have high-dimensional characteristics [2]. Benefiting from such high-dimensional characteristics, current applications of free-space optical (FSO) communication with OAM states are widely studied in the lab and real urban environments [3–10]. Naturally, the recognition of OAM modes in the receiving unit is one of the most important tasks for an optical communication system. A Gaussian mode converted by a forked hologram from the target OAM mode is the only mode that couples efficiently into the single-mode fiber [11]. More complicated computational holograms can be designed for the recognition of OAM superposition states [12]. Leach *et al.* presented the cascading additional Mach-Zehnder interferometers with dove prisms, which can sort OAM eigenstates into different paths [13]. By employing the Cartesian to log-polar transformation, one can convert the helically phased light beam corresponding to OAM state into a beam with a transverse phase gradient, and separate OAM eigenstates into different lateral positions [14–17]. Extensive research has also been carried out to characterize OAM modes by letting beams form

diffraction patterns by passing through well-designed masks, such as multipoint interference [18], triangular aperture diffraction [19], angular-double-slit interference [20], and gradually changing-period grating [21].

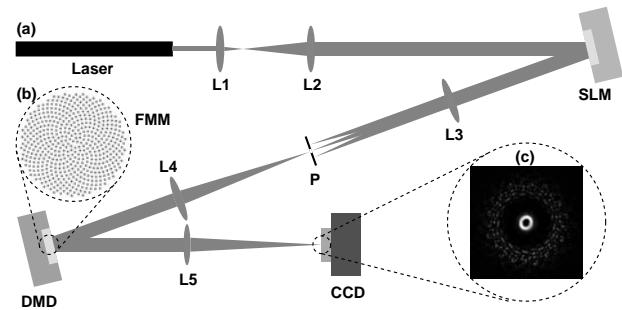
However, all the above OAM mode-recognizing methods require a complicated optical alignment process for FSO communication. Generally, the OAM of a light beam depends on the choice of the reference axis [22]. A pure OAM eigenstate will transform into the superposition of OAM states in a displaced coordinate frame [23] and result in the mixing of information between adjacent modes. In the standard approaches to FSO communication with the polarization of photons, the transmitting and receiving units with a shared reference frame are required. In 2012, Ambrosio *et al.* implemented the quantum communication [24] with hybrid polarization-OAM-entangled states, and this proposal is rotation-immune to the shared reference frame. Displacement of the reference frame also imposes serious obstacles to the application of FSO communication with OAM states. To overcome these obstacles, misalignment correction is implemented by using the mean square value of the OAM spectrum as an indicator [25]. However, OAM spectrum measurement with high precision under the case of misalignment is necessary before the OAM spectrum correction.

Recently, with the rapid increase in computing power, deep learning (DL) [26] has once again become a hot topic in various disciplines. Trained deep neural networks (DNNs) show state-of-the-art performance in imaging through scattering media [27–29], phase retrieval [30,31], structure light recognition [9,32–36], and creating new quantum experiments [37]. A milestone in the history of convolutional neural networks (CNNs) is the appearance of ResNet proposed by He *et al.* [38]. The core of the ResNet model is to establish shortcuts or skip connections between early layers and later layers, which helps in the backpropagation of the gradient during the training process, so as to train a deeper CNN. As a representative of CNNs, DenseNet [39] performs well in the ImageNet data set, it establishes dense connections between all early layers and later layers, and, specifically, each layer accepts all preceding layers as its additional inputs. Another major characteristic of DenseNet is the feature reuse through the connection of features on the channel. These characteristics allow DenseNet to achieve better performance than ResNet with fewer parameters and computational costs.

In this work, we implement an alignment-free fractal multipoint interferometer for hyperfine OAM mode recognition assisted by DL. By using a well-designed fractal multipoint mask (FMM) to sample the complex phase fronts of OAM modes, wealthy diffraction intensity patterns can be recorded for different OAM modes. Meanwhile, the diffraction patterns are stable against reference misalignment because of the inherent periodic structure of the FMM. Stochastic disturbances of three different parameters of the OAM states are set in the experiments: (i) beam waist size  $\omega \in [0.45, 0.55]$  mm; (ii) initial phase of OAM states  $\varphi_0 \in [0, 2\pi]$ ; (iii) lateral translation range along the  $x$  and  $y$  directions  $\Delta x, \Delta y \in [-0.25, 0.25]$  mm. Here, the maximum lateral displacement of the FMM we added with respect to the perfectly on-axis beam is about  $\pm 0.5$  beam waist size along the  $x$  and  $y$  directions, respectively. With the above three parameters changing randomly at the same time, we implement the recognition of OAM eigenstates with an accuracy of 100%. Adjacent OAM superposition states with a Bures distance (BD) close to 0.01 are also recognized with an accuracy higher than 98.3%. Benefiting from the simple FMM configuration and superhigh resolution recognition with high accuracy, our detection method is very useful for systems where the optical vortices are expected to be on very large scales, such as in FSO communication [9] and astronomical optical vortices [40].

## 2. METHODS

The experimental setup is shown in Fig. 1(a). A He-Ne laser with 633 nm wavelength is utilized as light source. After being collimated and expanded, the laser beam is projected on a phase-only spatial light modulator (SLM) (Hamamatsu model X10468) to generate the desired Laguerre-Gaussian (LG) modes. Then the LG modes are imaged on the plane of a digital micromirror device (DMD) (DLP4500, 1140 × 912 diamond pixel array of 7.6 μm × 7.6 μm mirrors) by a 4f imaging system L3 and L4. A well-designed FMM as shown in the inset of Fig. 1(b) is loaded on the DMD. After leaving the DMD plane, the beam passes through lens L5, and the far-field diffraction



**Fig. 1.** (a) Alignment-free fractal multipoint interferometer. Laser, He-Ne laser with 633 nm wavelength; L1, 50 mm lens; L2, 500 mm lens; SLM, phase-only spatial light modulator; L3, 300 mm lens; P, pinhole; L4, 300 mm lens; DMD, digital micromirror device; L5, 250 mm lens; CCD, charge-coupled device. (b) proposed FMM; (c) example of the far-field intensity patterns.

intensity pattern is collected by a charge-coupled device (CCD) (Lumenera INFINITY3-1C) with 1392 × 1040 pixels. We set the CCD to operate in 12-bit mode. Figure 1(c) shows an example of the recorded diffraction intensity patterns.

The LG modes have a complex field amplitude given by

$$LG_p^\ell(r, \varphi) \propto r^\ell L_p^\ell\left(\frac{2r^2}{\omega^2}\right) \exp\left(-\frac{r^2}{\omega^2}\right) \exp(-i\ell\varphi), \quad (1)$$

where  $(r, \varphi)$  are the radial and azimuthal coordinates, respectively.  $\omega$  is the beam waist, and  $L_p^\ell(2r^2/\omega^2)$  is the associated Laguerre polynomial.  $\ell$  is the topological charge, and  $p$  is the radial mode index. The complex amplitude field after the FMM is

$$U(x, y) = \sum_{n=1}^N \text{circ}\left[\frac{\sqrt{(x - x_n)^2 + (y - y_n)^2}}{r_0}\right] LG_p^\ell(r, \varphi), \quad (2)$$

where  $\text{circ}(x, y)$  is the transmittance function of the aperture in the FMM, and  $r_0$  is the radius of the FMM aperture.  $x_n, y_n$  are the central coordinates of the  $n$ th aperture. In the experiments, we implement the lateral translation misalignment by adding random lateral displacement  $\Delta x$  and  $\Delta y$  along the  $x$  and  $y$  directions, respectively, for the FMM. Here we employ a model for the pattern of florets in the head of a sunflower proposed by Helmut Vogel [41] in 1979 to arrange the position of each aperture, which is given by

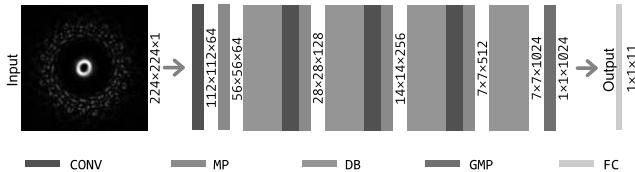
$$x_n = C_1 \sqrt{n} \cos\left(\frac{2\pi}{C_2 n}\right), \quad y_n = C_1 \sqrt{n} \sin\left(\frac{2\pi}{C_2 n}\right), \quad (3)$$

where  $C_1$  is the constant scaling factor, and  $C_2$  is the divergence angle. Considering the Fraunhofer limit, the far-field intensity pattern in the detector plane  $I$  is given by the Fourier transform of the field in the FMM plane,

$$I \propto |\mathcal{F}[U(x, y)]|^2, \quad (4)$$

and  $\mathcal{F}[\cdot]$  represents the Fourier transform.

To estimate the topological charge of the LG modes with the recorded intensity patterns  $I$ , we define  $I = H(\ell)$ , where the  $H(\cdot)$  represents the forward physical process that produces the diffraction pattern from the incident LG mode with the topological charge  $\ell$ . The optimization problem can be implicitly written as



**Fig. 2.** Schematic diagram of DenseNet-121. CONV, convolution layer; MP, max pooling layer; DB, dense block; GMP, global max pooling layer; FC, fully connected layer.

$$\hat{\ell} = \operatorname{argmin}_{\ell} \mathcal{L}\{H(\ell), I, R(\ell)\}, \quad (5)$$

where  $\hat{\ell}$  is the estimate of the inverse,  $\mathcal{L}$  is the objective function to minimize, and  $R(\ell)$  is the regularizer, or prior knowledge term that imposes constraints on the solution.

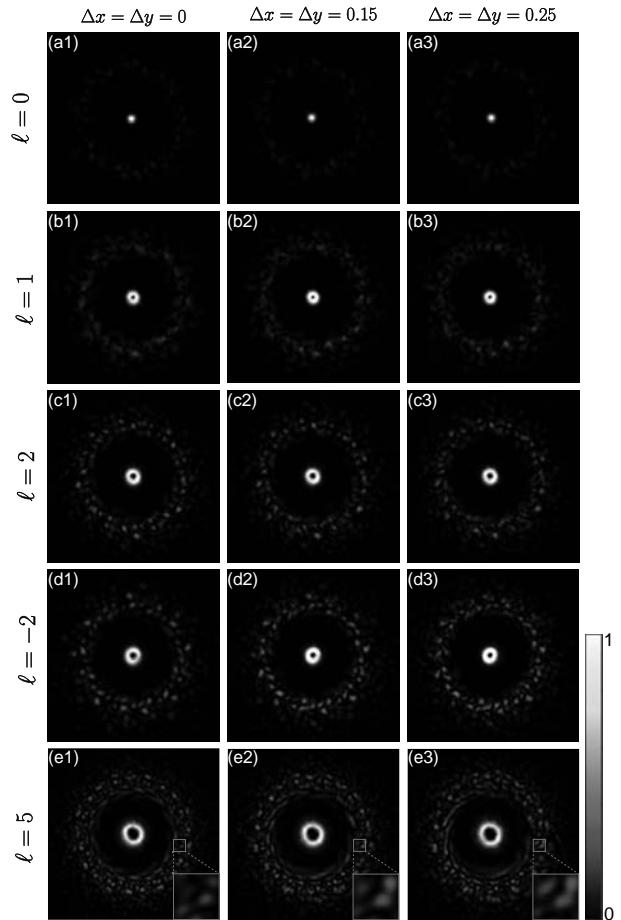
Here we adopt an end-to-end DL method, the DenseNet-121, to solve the above optimization problem. The architecture of the DenseNet-121 is shown in Fig. 2. The diffraction intensity patterns collected by the CCD are cropped and resized to  $224 \times 224$  pixels from raw  $1392 \times 1040$  pixels, with their corresponding state labels artificially added. The diffraction patterns and labels are paired to form the training set as inputs to the DenseNet-121. Then the inputs are fed to a convolution layer with filter size  $7 \times 7$  and strides (2, 2) followed by a  $3 \times 3$  max pooling layer with stride (2, 2). After that, four dense blocks with 6, 12, 24, and 16 convolution units, respectively, are implemented, and a  $1 \times 1$  convolution layer followed by  $2 \times 2$  max pooling is utilized as transition layers between two contiguous dense blocks. At the end of the last dense block, a global max pooling is performed; then a fully connected layer with a softmax classifier is attached. In DenseNet-121, rectified linear units (Relus) are used as the activation function, categorical-cross-entropy loss as loss function, and stochastic gradient descent (SGD) as optimizer. It is worth pointing out that the difference between the adjacent superposition states we selected is very small, and the inherent average pooling layer of DenseNet-121 will further weaken this difference. Therefore, we choose a max pooling layer to amplify small differences of the diffraction intensity patterns between different states. The program in our experiment was implemented on the Keras framework with Python 3.5, and sped up by a pair of GPUs (NVIDIA GTX 1080ti).

### 3. RESULTS

We first perform the DenseNet-121 to recognize LG eigenstates with topological charge  $\ell \in \{-5, -4, \dots, 5\}$  and  $p = 0$ . In order to test the robustness of the proposed method, stochastic disturbance of three parameters of the OAM states is set simultaneously for the acquisition of each diffraction intensity pattern: (i) beam waist size  $\omega \in [0.45, 0.55]$  mm; (ii) initial phase of OAM states  $\varphi_0 \in [0, 2\pi]$ ; (iii) lateral translation range along the  $x$  and  $y$  directions  $\Delta x, \Delta y \in [-0.25, 0.25]$  mm. A total of 1100 experimental diffraction intensity patterns and their corresponding topological charge  $\ell$  as labels are used as the data set, with 100 samples for each topological charge  $\ell$ . All 1100 samples are randomly shuffled, of

which the first 850 samples are used as the training set; the remaining 250 samples never participate in the training process.

Figure 3 shows the examples of recorded diffraction intensity patterns for LG eigenstates with topological charge  $\ell \in \{0, 1, \pm 2, 5\}$ . Figures 3(a1)–3(e1) are diffraction patterns when LG beams are perfectly on-axis. Figures 3(a2)–3(e2) and 3(a3)–3(e3) are diffraction patterns when lateral translations of  $\Delta x = \Delta y = 0.15$  mm and 0.25 mm are added on the LG beams, respectively. We collect each diffraction pattern in Fig. 3 by keeping the parameters  $\omega$  and  $\varphi_0$  to vary randomly according to the range mentioned above. The diffraction patterns show characters of donut-shaped intensity profile in the center with surrounding asteroid-belt-like speckles. The patterns in the center keep an intensity profile similar to the input LG eigenstates. The additional asteroid-belt-like speckles behave as completely different intensity profiles resulting from the interference of different spiral wavefronts, and it is helpful for the recognition of LG modes. Since the apertures of the proposed FMM are uniformly distributed from the center to



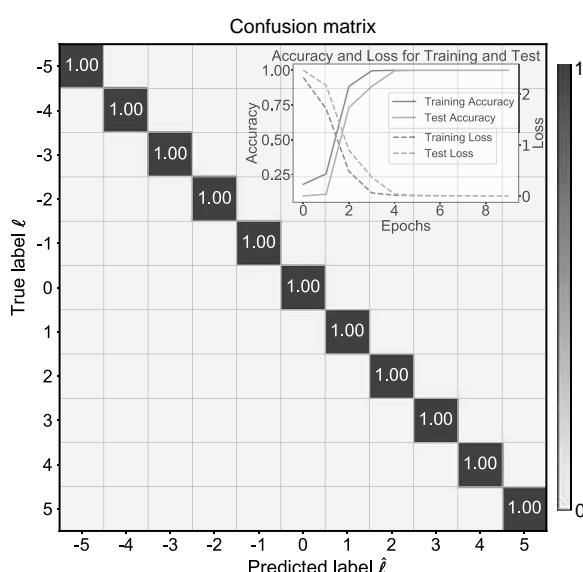
**Fig. 3.** Examples of the experimental diffraction intensity patterns for LG eigenstates with topological charge  $\ell \in \{0, 1, \pm 2, 5\}$  and different FMM displacements  $\Delta x = \Delta y \in \{0, 0.15, 0.25\}$  mm. In addition, all the diffraction patterns are obtained with stochastic disturbances of the other two parameters: (i) beam waist size  $\omega \in [0.45, 0.55]$  mm; (ii) initial phase of OAM states  $\varphi_0 \in [0, 2\pi]$ . Insets in (e1)–(e3) show the detailed profiles of the recorded intensity patterns.

the edge according to Eq. (3), it may help to maintain the details of the diffraction patterns to some extent. One can find that the enlarged details of the insets in Figs. 3(e1)–3(e3) are basically retained under different FMM displacements. Furthermore, this distribution of the proposed FMM is asymmetric with respect to any axis in its plane, which contributes to the robustness of distinguishing the diffraction patterns from positive and negative  $\ell$ . After careful tuning, the trained DenseNet-121 is fed with test data to evaluate its ability for classification. A normalized confusion matrix for misaligned LG eigenstates  $\ell \in \{-5, -4, \dots, 5\}$  and  $p = 0$  is shown in Fig. 4. All test data are correctly recognized, and an accuracy rate of 100% is achieved. As can be seen from the curves on the top right of Fig. 4, it only took five epochs for the DenseNet-121 to reach 100% accuracy in both the training and test sets. This indicates that the diffraction patterns formed by the LG eigenstates with different topological charge  $\ell$  are quite different and can be easily learned by the DenseNet-121 to achieve accurate classification.

To further demonstrate the performance of the proposed method, we experimentally implement the recognition of hyperfine OAM superposition states. We choose two mutually orthogonal bases,  $|\ell = \pm 1\rangle$ , to construct the Bloch sphere, and each point on the sphere is a superposition state constructed by this set of bases. As shown in Fig. 5(a), the state  $|\psi\rangle$  represented by an arbitrary point on the Bloch sphere is given by

$$|\psi\rangle = \cos \frac{\theta}{2} |1\rangle + \sin \frac{\theta}{2} e^{i\phi} |-1\rangle, \quad (6)$$

where  $\theta$  is the polar angle and  $\phi$  is the azimuthal angle. We take 200 and 400 points uniformly from  $[0, \pi]$  and  $[0, 2\pi]$  according to the respective ranges of  $\theta$  and  $\phi$ . In this way, 80,000 points with an interval of  $0.005\pi$  evenly distributed on the spherical surface are obtained, as schematically shown in Fig. 5(b). Without loss of generality, we randomly select an area on



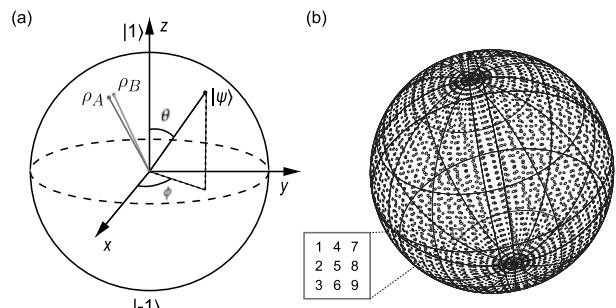
**Fig. 4.** Confusion matrix for the recognition of misaligned LG eigenstates  $\ell \in \{-5, -4, \dots, 5\}$  and  $p = 0$ , with the curves of accuracy and loss as functions of epochs on the top right.

the Bloch sphere with nine superposition states. The nine superposition states combined by  $\theta$  and  $\phi$  are shown in Table 1, and we name them Modes 1–9 for convenience. The values of  $\theta$  are  $0.52\pi$ ,  $0.525\pi$ , and  $0.53\pi$ , while the values of  $\phi$  are  $0.02\pi$ ,  $0.025\pi$ , and  $0.03\pi$ . The red box in Fig. 5(b) schematically shows the position distribution of the nine superposition states on the Bloch sphere. We utilize the BD [42]  $D_B^2[\rho_A, \rho_B] = 2[1 - \sqrt{F(\rho_A, \rho_B)}]$ , where  $F(\rho_A, \rho_B) = [\text{Tr} \sqrt{\sqrt{\rho_A} \rho_B \sqrt{\rho_A}}]^2$  is the fidelity of the two states, to calculate the distance between adjacent modes  $\rho_A$  and  $\rho_B$  on the Bloch sphere. The distance between two horizontal (e.g., Modes 4 and 7) or vertical (e.g., Modes 1 and 2) adjacent modes of the nine modes we selected on the Bloch sphere is 0.01, while the distance between diagonally positioned adjacent modes (e.g., Modes 5 and 9) is 0.015. This means that any two adjacent modes are so similar that their corresponding diffraction patterns are also very similar and difficult to recognize.

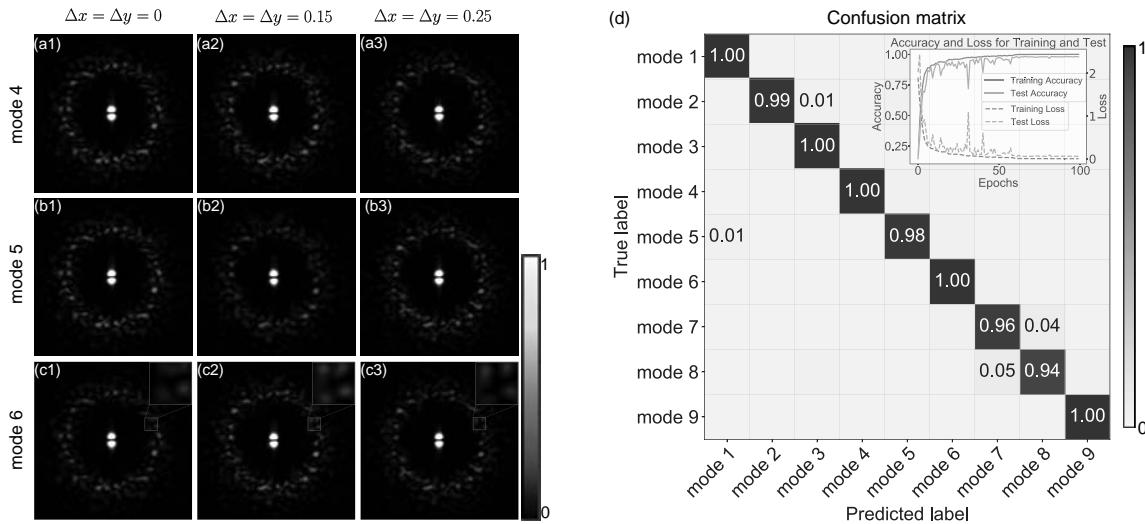
In the experiments, stochastic disturbances of parameters, (i) beam waist size  $\omega \in [0.45, 0.55]$  mm, (ii) initial phase  $\varphi_0 \in [0, 2\pi]$ , and (iii) lateral translation  $\Delta x, \Delta y \in [-0.25, 0.25]$  mm, are also introduced in these superposition states. A total of 9000 recorded diffraction intensity patterns and their corresponding labels are used as the data set: 1000 samples for each category, 7000 of which are used as the training set and 2000 are used as the test set. The diffraction intensity patterns for superposition states of Mode 4, Mode 5, and Mode 6 under different stochastic disturbances of the above three parameters are shown in Figs. 6(a1)–6(c3). Similar to the results of eigenstates, the diffraction patterns for superposition states also consist of two parts: central lobes that have the similar intensity profile to the input superposition states and the surrounding asteroid-belt-like speckles. One can see that the diffraction patterns formed by different superposition states are highly

**Table 1.** Nine Superposition States Combined by  $\theta$  and  $\phi$

	$\phi = 0.020\pi$	$\phi = 0.025\pi$	$\phi = 0.030\pi$
$\theta = 0.520\pi$	Mode 1	Mode 4	Mode 7
$\theta = 0.525\pi$	Mode 2	Mode 5	Mode 8
$\theta = 0.530\pi$	Mode 3	Mode 6	Mode 9



**Fig. 5.** Schematic diagram of a Bloch sphere constructed with  $|\ell = \pm 1\rangle$  bases. (a) An arbitrary state  $|\psi\rangle$  on the sphere and two adjacent superposition states  $\rho_A$  and  $\rho_B$ ; (b) sphere is divided into 80,000 points corresponding to 80,000 states; the red box schematically indicates the position distribution of the nine selected superposition states for the experiments.



**Fig. 6.** Experimental results of hyperfine LG superposition states. (a)–(c) Examples of the recorded diffraction intensity patterns for OAM superposition states under different misaligned configurations. The collection of each diffraction pattern in the figure is carried out with stochastic disturbances of the other two parameters: (i) beam waist size  $\omega \in [0.45, 0.55]$  mm; (ii) initial phase of OAM states  $\varphi_0 \in [0, 2\pi]$ . To show the image clearly, we reduce the contrast of the image by setting the values of  $I = 0.3 \times \max\{I\}$ , where  $I > 0.3 \times \max\{I\}$ ; (d) confusion matrix for superposition states from Mode 1 to Mode 9 with the curves of accuracy and loss as functions of epochs on the top right.

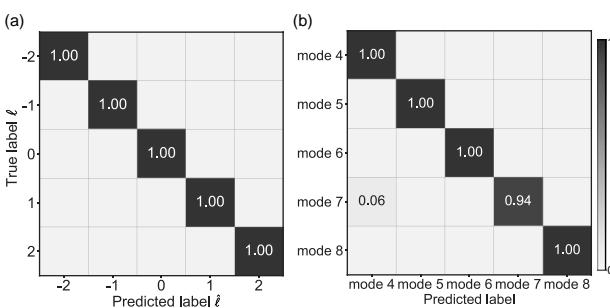
semblable. However, the well-tuned DenseNet-121 achieved an accuracy of 98.35%, and only a small number of adjacent states are misjudged, as shown in Fig. 6(d). The recognition accuracy of superposition states is lower than that of eigenstates, but it is still improving. Compared with eigenstates, the training of superposition states is obviously more time-consuming, as shown in the curves on the top right of Fig. 6(d). This is because the superposition states require more training samples and the similarity between the superposition states is large, which makes it difficult for them to converge. On the one hand, the disturbance of the experimental device or the environment is likely to have a huge impact on the experimental results when the adjacent superposition states are highly similar. On the other hand, the roughly retained details in the insets of Figs. 6(c1)–6(c3) and the raise in the number of training data help DenseNet-121 achieve better feature extraction. In fact, if one reduces the intensity of the illumination laser beam, the asteroid-belt-like speckles will be drowning in the electron

noise of the camera. The predictive accuracy of the hyperfine OAM states is reduced as we lose the detailed information of the asteroid-belt-like speckles.

Here, to fully demonstrate the robustness of the proposed method, we implement the hyperfine OAM mode recognition for the case of  $p \neq 0$ . In experiments, we set  $p = 1$ , and all other conditions remain the same as the above experiments. The normalized partial confusion matrices of the eigenstates  $\ell \in \{-2, -1, \dots, 2\}$  and superposition states of Mode 4 to Mode 8 in Table 1 are shown in Figs. 7(a) and 7(b), respectively. Similar to the result in the case of  $p = 0$ , the DenseNet-121 easily realized an accuracy of 100% for the eigenstates, and an accuracy of 98.82% for superposition states with the radial mode index  $p = 1$ .

#### 4. CONCLUSION

In conclusion, we experimentally implemented hyperfine OAM mode recognition under strong misalignment by using an alignment-free fractal multipoint interferometer assisted by DL. The misalignment includes three stochastic disturbances of parameters: (i) beam waist size  $\omega \in [0.45, 0.55]$  mm; (ii) initial phase  $\varphi_0 \in [0, 2\pi]$ ; (iii) lateral translation  $\Delta x, \Delta y \in [-0.25, 0.25]$  mm. Here, the maximum lateral misalignment of the FMM we added with respect to the perfectly on-axis beam is about  $\pm 0.5$  beam waist size along the  $x$  and  $y$  directions, respectively. The well-tuned DenseNet-121 is demonstrated to be robust for recognizing very similar superposition states with a small BD of 0.01 between adjacent modes under the above strong misalignment. Benefiting from the robustness of the proposed method and simple FMM configuration, this scheme shows potential application for FSO communication where the optical vortices are expected to be on a large scale and the misalignment between the transmitting and receiving units is inevitable.



**Fig. 7.** Confusion matrix of LG modes with  $p = 1$ . (a) Confusion matrix for LG eigenstates of  $\ell \in \{-2, -1, \dots, 2\}$  and  $p = 1$ ; (b) confusion matrix for superposition states from Mode 4 to Mode 8. Mode 4 to Mode 8 here represent the superposition states with  $p = 1$  at the same positions on the Bloch sphere in the case of  $p = 0$ .

**Funding.** National Natural Science Foundation of China (11534008, 91736104); Ministry of Science and Technology of the People's Republic of China (2016YFA0301404); Fundamental Research Funds for the Central Universities.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

1. L. Allen, M. W. Beijersbergen, R. Spreeuw, and J. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A* **45**, 8185–8189 (1992).
2. A. M. Yao and M. J. Padgett, "Orbital angular momentum: origins, behavior and applications," *Adv. Opt. Photon.* **3**, 161–204 (2011).
3. A. Vaziri, G. Weihs, and A. Zeilinger, "Experimental two-photon, three-dimensional entanglement for quantum communication," *Phys. Rev. Lett.* **89**, 240401 (2002).
4. G. Gibson, J. Courtial, M. J. Padgett, M. Vasnetsov, V. Pas'ko, S. M. Barnett, and S. Franke-Arnold, "Free-space information transfer using light beams carrying orbital angular momentum," *Opt. Express* **12**, 5448–5456 (2004).
5. J. Wang, J.-Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, and A. E. Willner, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nat. Photonics* **6**, 488–496 (2012).
6. N. Bozinovic, Y. Yue, Y. Ren, M. Tur, P. Kristensen, H. Huang, A. E. Willner, and S. Ramachandran, "Terabit-scale orbital angular momentum mode division multiplexing in fibers," *Science* **340**, 1545–1548 (2013).
7. M. Krenn, R. Fickler, M. Fink, J. Handsteiner, M. Malik, T. Scheidl, R. Ursin, and A. Zeilinger, "Communication with spatially modulated light through turbulent air across Vienna," *New J. Phys.* **16**, 113028 (2014).
8. A. E. Willner, H. Huang, Y. Yan, Y. Ren, N. Ahmed, G. Xie, C. Bao, L. Li, Y. Cao, Z. Zhao, J. Wang, M. P. J. Lavery, M. Tur, S. Ramachandran, A. F. Molisch, N. Ashrafi, and S. Ashrafi, "Optical communications using orbital angular momentum beams," *Adv. Opt. Photon.* **7**, 66–106 (2015).
9. M. Krenn, J. Handsteiner, M. Fink, R. Fickler, R. Ursin, M. Malik, and A. Zeilinger, "Twisted light transmission over 143 km," *Proc. Natl. Acad. Sci. USA* **113**, 13648–13653 (2016).
10. M. P. Lavery, C. Peuntinger, K. Günthner, P. Banzer, D. Elser, R. W. Boyd, M. J. Padgett, C. Marquardt, and G. Leuchs, "Free-space propagation of high-dimensional structured optical fields in an urban environment," *Sci. Adv.* **3**, e1700552 (2017).
11. A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the orbital angular momentum states of photons," *Nature* **412**, 313–316 (2001).
12. E. Bolduc, N. Bent, E. Santamato, E. Karimi, and R. W. Boyd, "Exact solution to simultaneous intensity and phase encryption with a single phase-only hologram," *Opt. Lett.* **38**, 3546–3549 (2013).
13. J. Leach, M. J. Padgett, S. M. Barnett, S. Franke-Arnold, and J. Courtial, "Measuring the orbital angular momentum of a single photon," *Phys. Rev. Lett.* **88**, 257901 (2002).
14. G. C. Berkhout, M. P. Lavery, J. Courtial, M. W. Beijersbergen, and M. J. Padgett, "Efficient sorting of orbital angular momentum states of light," *Phys. Rev. Lett.* **105**, 153601 (2010).
15. M. Mirhosseini, M. Malik, Z. Shi, and R. W. Boyd, "Efficient separation of the orbital angular momentum eigenstates of light," *Nat. Commun.* **4**, 2781 (2013).
16. Y. Wen, I. Chremmos, Y. Chen, J. Zhu, Y. Zhang, and S. Yu, "Spiral transformation for high-resolution and efficient sorting of optical vortex modes," *Phys. Rev. Lett.* **120**, 193904 (2018).
17. Y. Wen, I. Chremmos, Y. Chen, G. Zhu, J. Zhang, J. Zhu, Y. Zhang, J. Liu, and S. Yu, "Compact and high-performance vortex mode sorter for multi-dimensional multiplexed fiber communication systems," *Optica* **7**, 254–262 (2020).
18. G. C. Berkhout and M. W. Beijersbergen, "Method for probing the orbital angular momentum of optical vortices in electromagnetic waves from astronomical objects," *Phys. Rev. Lett.* **101**, 100801 (2008).
19. J. Hickmann, E. Fonseca, W. Soares, and S. Chávez-Cerda, "Unveiling a truncated optical lattice associated with a triangular aperture using light's orbital angular momentum," *Phys. Rev. Lett.* **105**, 053904 (2010).
20. R. Liu, J. Long, F. Wang, Y. Wang, P. Zhang, H. Gao, and F. Li, "Characterizing the phase profile of a vortex beam with angular-double-slit interference," *J. Opt.* **15**, 125712 (2013).
21. K. Dai, C. Gao, L. Zhong, Q. Na, and Q. Wang, "Measuring OAM states of light beams with gradually-changing-period gratings," *Opt. Lett.* **40**, 562–565 (2015).
22. G. Molina-Terriza, J. P. Torres, and L. Torner, "Management of the angular momentum of light: preparation of photons in multidimensional vector states of angular momentum," *Phys. Rev. Lett.* **88**, 013601 (2001).
23. M. Vasnetsov, V. Pas'ko, and M. Soskin, "Analysis of orbital angular momentum of a misaligned optical beam," *New J. Phys.* **7**, 46 (2005).
24. V. D'Ambrosio, E. Nagali, S. P. Walborn, L. Aolita, S. Slussarenko, L. Marrucci, and F. Sciarrino, "Complete experimental toolbox for alignment-free quantum communication," *Nat. Commun.* **3**, 961 (2012).
25. Y.-D. Liu, C. Gao, X. Qi, and H. Weber, "Orbital angular momentum (OAM) spectrum correction in free space optical communication," *Opt. Express* **16**, 7091–7101 (2008).
26. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
27. M. Lyu, H. Wang, G. Li, S. Zheng, and G. Situ, "Learning-based lensless imaging through optically thick scattering media," *Adv. Photon.* **1**, 036002 (2019).
28. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–1190 (2018).
29. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**, 16032–16046 (2019).
30. M. Deng, S. Li, A. Goy, I. Kang, and G. Barbastathis, "Learning to synthesize: robust phase retrieval at low photon counts," *Light Sci. Appl.* **9**, 1 (2020).
31. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**, 243902 (2018).
32. Q. Zhao, S. Hao, Y. Wang, L. Wang, X. Wan, and C. Xu, "Mode detection of misaligned orbital angular momentum beams based on convolutional neural network," *Appl. Opt.* **57**, 10152–10158 (2018).
33. Z. Liu, S. Yan, H. Liu, and X. Chen, "Superhigh-resolution recognition of optical vortex modes assisted by a deep-learning method," *Phys. Rev. Lett.* **123**, 183902 (2019).
34. P. Wang, J. Liu, L. Sheng, Y. He, W. Xiong, Z. Huang, X. Zhou, Y. Li, S. Chen, X. Zhang, and D. Fan, "Convolutional neural network-assisted optical orbital angular momentum recognition and communication," *IEEE Access* **7**, 162025 (2019).
35. S. Sharifi, Y. Banadaki, G. Veronis, and J. P. Dowling, "Towards classification of experimental Laguerre-Gaussian modes using convolutional neural networks," *Opt. Eng.* **59**, 076113 (2020).
36. T. Giordani, A. Suprano, E. Polino, F. Acanfora, L. Innocenti, A. Ferraro, M. Paternostro, N. Spagnolo, and F. Sciarrino, "Machine learning-based classification of vector vortex beams," *Phys. Rev. Lett.* **124**, 160401 (2020).
37. A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, "Active learning machine learns to create new quantum experiments," *Proc. Natl. Acad. Sci. USA* **115**, 1221–1226 (2018).
38. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
39. G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4700–4708.
40. M. Harwit, "Photon orbital angular momentum in astrophysics," *Astrophys. J.* **597**, 1266–1270 (2003).
41. H. Vogel, "A better way to construct the sunflower head," *Math. Biosci.* **44**, 179–189 (1979).
42. J. Dajka, J. Łuczka, and P. Hänggi, "Distance between quantum states in the presence of initial qubit-environment correlations: a comparative study," *Phys. Rev. A* **84**, 032120 (2011).

# PHOTONICS Research

## Experimental study of neuromorphic node based on a multiwaveband emitting two-section quantum dot laser

GEORGE SARANTOGLOU,<sup>1,3</sup> MENELAOS SKONTRANIS,<sup>1</sup> ADONIS BOGRIS,<sup>2</sup> AND CHARIS MESARITAKIS<sup>1,\*</sup> 

<sup>1</sup>Department of Information and Communication Systems Engineering, University of the Aegean, Palama 2, Karlovassi 83200 Samos, Greece

<sup>2</sup>Department of Informatics and Computer Engineering, University of West Attica, Aghiou Spiridonos, 12243 Egaleo, Athens, Greece

<sup>3</sup>e-mail: gsarantoglou@aegean.gr

\*Corresponding author: cmesar@aegean.gr

Received 27 October 2020; revised 18 January 2021; accepted 25 January 2021; posted 25 January 2021 (Doc. ID 413371); published 16 March 2021

In this work, we present experimental results concerning excitability in a multiband emitting quantum-dot-based photonic neuron. The experimental investigation revealed that the same two-section quantum dot laser can be tuned through a simple bias adjustment to operate either as a leaky integrate and fire or as a resonate and fire neuron. Furthermore, by exploiting the inherent multiband emission of quantum-dot devices revealed by the existence of multiple lasing thresholds, a significant enhancement in the neurocomputational capabilities, such as spiking duration and firing rate, is observed. Spike firing rate increased by an order of magnitude that leads to an enhancement in processing speed and, more importantly, neural spike duration was suppressed to the picosecond scale, which corresponds to a significant temporal resolution enhancement. These new regimes of operation, when combined with thermal insensitivity, silicon cointegration capability, and the fact that these multiband mechanisms are also present in miniaturized quantum-dot devices, render these neuromorphic nodes a proliferating platform for large-scale photonic spiking neural networks. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.413371>

### 1. INTRODUCTION

Bio-inspired computing has emerged in the last five years as a promising candidate to tackle the inefficiency of conventional von Neumann architectures to address machine-learning related problems [1]. More specifically, the vision of dedicated hardware that mimics the principle of operation of biological neural circuits is extremely appealing performance- and energy-consumption-wise. Having in mind these two aspects, photonic technology offers a proliferating platform for the development of such systems due to advantages such as low loss, inherent parallelism through wavelength-time multiplexing, ultralow response time, and similarity of dynamics with biological neurons [1]. Semiconductor lasers are active photonic components that fit really well in the neuromorphic landscape because they possess all the aforementioned advantages. In previous works, laser neurons have been examined both theoretically and experimentally utilizing different schemes, namely, coherent/noncoherent optical injection [2–4], optical feedback [5], two-section lasers [gain and saturable absorber (SA)] [6–8],

microdisk lasers [9], 2D photonic crystals [10], and polarization-switching vertical cavity surface emitting lasers (VCSELs) [11].

With respect to photonic-based solutions, quantum dot (QD) lasers in particular, can be utilized as extremely efficient neuromorphic nodes, since they present multiple benefits compared to traditional quantum well (QW) devices, such as temperature insensitivity and stability [12–14]. These aspects are of utmost importance for successful large-scale integration of laser-based spiking neural networks. Additionally, QDs can make silicon-based lasers possible [15], paving the way to seamless integration with other silicon photonic neuromorphic architectures [16]. In the literature, single-section QD laser neurons in optical injection schemes have been studied both theoretically and experimentally [17–19]. Moreover, we have recently provided experimental evidence of neuromorphic operation from a QD two-section device, at the same time achieving all-optical isomorphism to excitatory and inhibitory neurons [20].

In order to harness the full potential of biological neural motifs, it is important to develop photonic components able to replicate the neurocomputational characteristics of biological neurons. A detailed picture of the most important neurocomputational properties is presented in Table 1 [21]. The most commonly encountered neurons are Class I integrators and resonators. These two types of neurons differ in terms of the existence of fast subthreshold small amplitude oscillations that can be unmasked by the presence of noise. In particular, integrators reveal no subthreshold oscillations because they are highly damped systems, while resonators do reveal such oscillations. The other neuronal differences that exist between these two categories stem from this basic difference [21]. In particular, integrators reveal an adjustable firing rate depending on the input strength (rate encoding scheme). Moreover, they present a well-defined neural threshold, which means that when the external triggering is below a certain value, no spikes are observed, whereas beyond that threshold, spike events of relatively constant amplitude are generated (all-or-none spiking). On the other hand, in resonators, the firing rate is limited in a small frequency range that is dictated by the subthreshold oscillations. In general, they do not exhibit a well-defined threshold value, but a threshold area. When the energy of external triggering is within a certain energy area, spike events can be produced whose amplitude is variable, depending on the strength of the applied triggering [21]. This results in the absence of all-or-nothing spiking, which in experiments can be observed through spike events whose amplitude presents large standard deviation due to the presence of noise-induced triggering. An important exception is the case of slow-fast systems, where the threshold area is so thin that it resembles a well-defined threshold known as the quasi-threshold [22]. In this case, the transition from resting to spiking dynamics is known as a Canard explosion [23], a situation that has been observed in optically injected lasers [18,24] and photonic crystals [10]. Last but not least, resonators present frequency preference, which means that firing efficiency is enhanced when the frequency of the input signal matches the subthreshold oscillation frequency.

Application-wise, the integrators are important as processing units that are capable of encoding information in the time signature of various all-or-none spike events. Resonators, due to their frequency selectivity, can act as bandpass filters that transform the input signal in a spike train. Moreover, they can be used to achieve selective communication between different neural groups [25]. These two types of neurons can be efficiently transferred in the photonic domain, through QD devices, due to the aforementioned pivotal advantages that they

process. Up to now, QDs have been exploited so as to realize integrator dynamics in a two-section laser configuration [20] and through optical injection-based triggering [17]. With respect to resonators, the state of the art in QDs consists of an optically injected, edge-emitting, single-section laser, with emitting only from the ground-state (GS) waveband. This device allows the recording of microsecond spikes whose origin could be traced to intracavity thermal effects [18]. Thus, this implementation exhibits limited temporal resolution compared to conventional QW-based resonators [26]. This reduction in temporal resolution (spike duration) can greatly impact the accuracy of bio-inspired learning techniques such as spike-dependent plasticity [27], as well as the accuracy of the various spike encoding schemes [28].

In this work, our goal is to further advance the concept of two-section QD neurons introduced in our previous work [20]. Compared to Ref. [20], we provide an in-depth analysis of the excitatory QD neuron, which is the most common in neural networks, and present its manifestation under various laser bias conditions. In particular, we present evidence that the two-section QD laser operates as an integrator by exploiting the thermo-optic coupling, whereas the *Q*-switching effect that is strongly related to the multiwaveband-related dynamics is linked to resonate and fire operation. Moreover, the study of different bias regimes reveals for the first time the impact of multiwaveband emission, not on the neuron's type (excitatory/inhibitory) as it was shown in Ref. [20], but on its performance. The existence of multiwaveband dynamics allowed for spike events of subnanosecond temporal width, thus leading to highly improved temporal resolution and, additionally, to an increase of the firing rate by 1 order of magnitude. This improvement is more important, when considering the fact that the devices used in this work have an increased footprint and achieve performance similar to miniaturized lasers. Taking into consideration that the multiband behavior exploited here is also present in small footprint QD structures, this work paves new exploration paths for QD neurons. Furthermore, since a two-section device can be triggered both electro-optically through current modulation [29] and all-optically by applying external optical injection [30], we employ a twofold approach; on one hand we use the solitary laser neuron's injection current as a neural triggering mechanism, which is the simplest way to determine its neurocomputational properties [30]. Then we demonstrate an all-optical, application-focused scenario and analyze the effect of multiwaveband emission in the temporal resolution by triggering the laser with mode-locked pulses from a master QD laser. This twofold investigation signifies that

**Table 1. Major Neurocomputational Properties [21]**

Properties	Integrators		Resonators	
Excitability	Class I	No	Class II	Class II
Subthreshold oscillations		No		Yes
Frequency preference		No		Yes
Threshold		Well defined		May not be defined
All-or-none spiking		Yes		No (*yes)
Bistability	No	Yes	No	Yes
Bifurcation	Saddle node on limit cycle	Saddle node off limit cycle	Supercritical Hopf (*canard explosion)	Subcritical Hopf

the proposed scheme can provide excitability through both electro-optic and all-optical triggering.

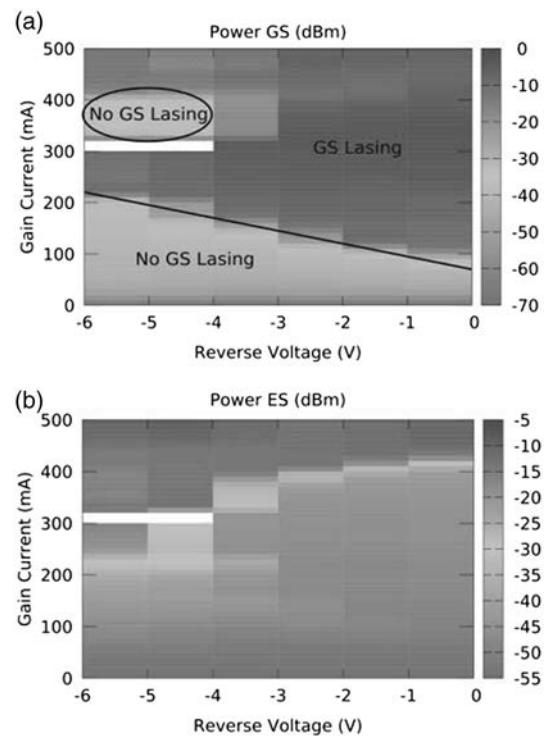
## 2. EXPERIMENTAL SETUP

The full-scale experimental setup, illustrated in Fig. 1, consists of two InAs/InGaAs Fabry–Perot, edge emitting, two-section QD lasers, connected in a master laser (ML)/slave laser (SL) configuration. The ML acts as the external stimulus by operating in a GS passive mode-locking regime, emitting pulses with a temporal width of  $<10$  ps. The active material consists of 10 QD layers into a 440 nm GaAs waveguide surrounded by an Al<sub>35%</sub>Ga<sub>65%</sub>As layer, whereas the cavity length is 4 mm. The SL acts as a photonic neuron. Its cavity length is 2 mm, and it consists of five QD layers. The gain/SA ratio is 0.85/0.15 for both lasers. The temperature is controlled by a Peltier element in a closed control loop. The temperatures are set to 33°C for the ML and 20°C for the SL. This temperature variation is set so their optical spectra partially overlap, despite the slightly different emission wavelengths of each structure.

At first, we study only the behavior of the SL under injection current variations; thus only a part of the aforementioned setup is used (Part I in Fig. 1), while later the full-scale unidirectional optical injection setup is exploited (Part II in Fig. 1). A variable optical attenuator (VOA) is used to control the optical power injected in the SL. A photodiode with 6 GHz bandwidth, followed by a real-time oscilloscope with 40 Gsample/s data acquisition capability is used to track the optical output of the SL, and an optical spectrum analyzer (OSA) with 0.05 nm resolution is used to measure the optical spectrum at both bands.

## 3. EXPERIMENTAL INVESTIGATION OF MULTIPLE GS THRESHOLDS

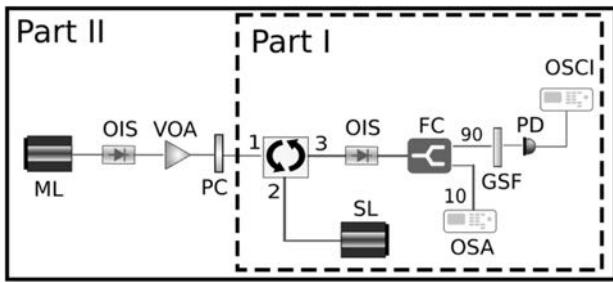
Excitability in a conventional two-section laser is achieved by biasing it below its lasing threshold and applying an external stimulus [30]. In this case, the SA acts as a threshold, determining the necessary strength of the input signal so as to trigger an optical spike [20]. In the case of QD-based lasers, the multi-waveband emission capabilities, namely, GS/excited state



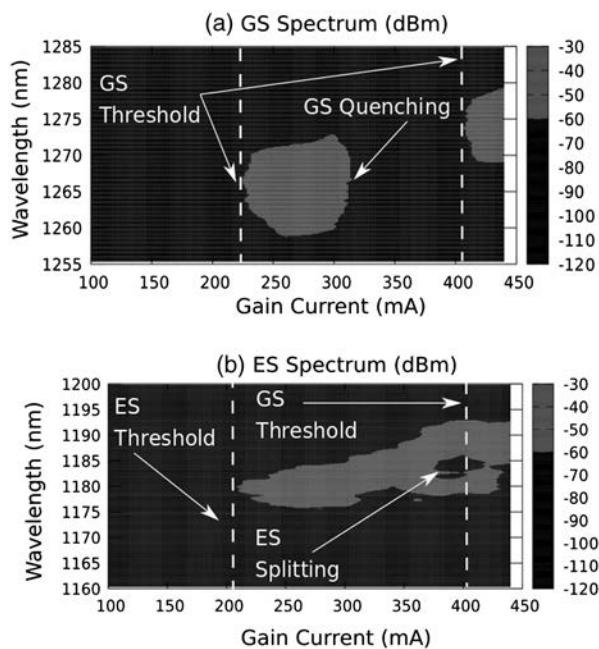
**Fig. 2.** (a) GS optical power of the QD SL, for various gain current and SA reverse voltage bias conditions; (b) ES optical power of the QD SL, for various gain current and SA reverse voltage bias conditions. The white regions refer to corrupt measurements, which, however, correspond to bias regimes that are not related to the neurocomputational properties of the SL.

(GS/ES) emission, allow multiple subthreshold regimes where emission occurs from a single waveband, while the other one is set below its lasing threshold. This behavior can affect the laser's dynamics and modify the anticipated excitable operation and performance. In this work, we focus on operational regimes where GS emission is below the threshold, while ES emission can be on or off. In this scope, in Fig. 2 the optical power for the GS [Fig. 2(a)] and ES band [Fig. 2(b)] of the solitary SL as a function of the applied gain current and reverse voltage at the SA is presented.

The lasing threshold of GS emission increases with the increase of reverse voltage due to the increase of the SA's unsaturated losses [31]. Moreover, GS lasing appears to have two discrete thresholds for reverse voltages above 4 V. The first is observed for low current injection ( $\approx 200$  mA) and the second one for higher current injection ( $\approx 400$  mA), associated with ES lasing [compare Figs. 2(a) and 2(b)]. In order to shed light on the origin of the two discrete GS thresholds, in Figs. 3(a) and 3(b), the GS and ES optical spectra are demonstrated versus injection current for 5 V reverse voltage. For low injection current, GS and ES lasing is achieved simultaneously due to the small cavity length and the low number of QD layers (first GS lasing threshold) [32,33]. As injection current increases, GS lasing is suppressed due to the GS quenching effect [34]. After this point, higher injection current leads to ES spectral splitting, where two discrete sets of longitudinal modes are



**Fig. 1.** Experimental setup. ML, master laser; SL, slave laser; VOA, variable optical attenuator; PC, polarization controller; OC, optical circulator; FC, fiber coupler; OSA, optical spectrum analyzer; GSF, ground-state passband filter; PD, photodiode; and OSCI, oscilloscope. The dashed box contains the first experiment (Part I), which was based on the electro-optic triggering of the SL through current injection, whereas the full line box contains the full unidirectional optical injection setup (Part II).



**Fig. 3.** Development of (a) GS and (b) ES spectrum with gain current for  $V_{\text{rev}} = 5$  V.

lasing. The spectral separation ( $\Delta\lambda$ ) between the centers of each set of modes increases with current and can extend up to  $\Delta\lambda = 9.5$  nm for 400 mA. When injection current increases beyond 400 mA,  $\Delta\lambda$  starts being reduced, and GS is forced again above the threshold (second GS threshold). For higher current injection (440 mA), ES splitting completely vanishes. This behavior is observed here for the first time and is probably linked to gain suppression mechanisms similar to those responsible for GS spectral splitting [35].

This spectral evolution can provide an insight into the dynamics responsible for the “second” GS threshold. In particular, ES splitting is linked to gain suppression effects like spectral hole burning at the central longitudinal modes [35,36]. Therefore, this gain suppression reduces the stimulated emission process at the ES, allowing the reaccumulation of free carriers at the GS, and subsequently evoking GS lasing. When GS lasing is triggered, GS free carriers are depleted, and the flow of carriers in the ES is reduced. This effect, in turn, reduces the gain in the ES and gradually terminates the saturation-induced splitting effect (see Fig. 3).

The key feature is that the carrier dynamics at each band in this new regime are inherently different compared to the conventional single-band threshold of devices. Thus, in the following paragraphs, we focus on analyzing the excitability for these two regimes, which involve subthreshold GS biasing.

#### 4. EXPERIMENTAL INVESTIGATION OF EXCITABILITY REGIMES USING INJECTION CURRENT

The neurocomputational properties of an excitable system can be deduced by performing a bifurcation analysis (see Table 1) [21]. In particular, in excitable systems, the variation

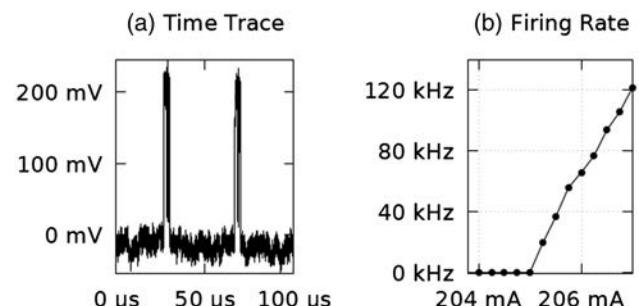
of a single parameter known as the bifurcation parameter leads to the transition from a resting state to oscillatory dynamics. In two-section devices, the simplest way to study the excitability-related bifurcation is the variation of the current injection close to the laser threshold [30].

##### A. Neurocomputational Properties Close to the Low (First) GS Current Threshold

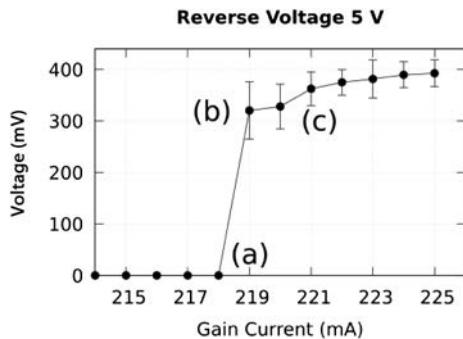
For reverse voltage  $V_{\text{rev}} \leq 4$  V, only the low-injection GS threshold exists. For  $V_{\text{rev}} = 4$  V, the transition from resting dynamics to self-pulsations is characterized by microsecond all-or-none spikes with very low repetition frequency [Fig. 4(a)], whose origin is related to the optothermal coupling effect [37,38]. Unlike single-section QD devices [38], no fast small-amplitude oscillations are observed at the resting state, which indicates that the SL behaves as an integrator. The firing rate increases with gain current rate encoding [Fig. 4(b)] as in Class I integrator systems (Table 1). This behavior may be linked to a homoclinic saddle node bifurcation, which was observed recently in single-section QD lasers [39]. It is important to note that this dynamic regime can be traced for all  $V_{\text{rev}} \leq 4$  V.

On the other hand, for  $V_{\text{rev}} = 5$  V and gain current close to 218 mA, although the laser is set below/near its threshold, perturbations that evoke a radical shift in the laser’s behavior are observed. In Fig. 3, it can be seen that for this bias regime GS/ES emission thresholds coincide, meaning that the GS dynamics are also affected by the ES band. Although the power of ES is weak in this case, as shown in Fig. 2, its dynamics still affect the GS waveband through the scattering rates as opposed to the case of 4.5 V, where the influence of the ES waveband becomes significant—close to 350 mA. The temporal width of the generated spikes is considerably reduced, from the microsecond to the nanosecond scale. This indicates that these pulsations are linked to  $Q$ -switching effects [29], which in turn are possibly linked to GS/ES interplay [40]. The mean amplitude and standard deviation of the spikes recorded at the photodiode are depicted in Fig. 5 as a function of the gain current. The low standard deviation indicates the property of all-or-none spiking.

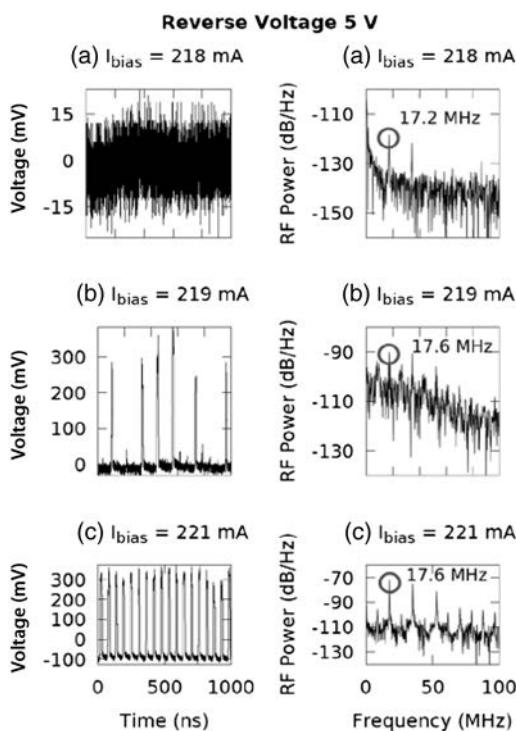
In Fig. 6, the RF spectrum and time traces for 218 mA (a), 219 mA (b), and 221 mA (c) are presented. For 218 mA, the laser is below its threshold, and no spikes are recorded, but in the RF spectrum, a clear peak at 17.2 MHz is evident, corresponding to subthreshold pulsations unmasked by noise [21]. For 219 mA, the resting state coexists with noise-triggered spike



**Fig. 4.** (a) Two opto-thermal spike events for  $V_{\text{rev}} = 4$  V and  $I_{\text{bias}} = 205.25$  mA; (b) firing rate of the generated spike events as a function of the gain current for the same reverse voltage bias.



**Fig. 5.** Spike amplitude for  $V_{\text{rev}} = 5$  V close to the low GS injection threshold. The points (a), (b), and (c) correspond to  $I_{\text{bias}} = 218$ , 219, and 221 mA.



**Fig. 6.** RF spectra and time traces for 5 V reverse voltage and  $I_{\text{bias}} = 218$  (a), 219 (b), and 221 mA (c).

events. Interestingly, these spikes are built upon the 17.6 MHz subthreshold oscillations, as the RF spectrum reveals. By further increasing the gain current to 221 mA, the laser fully enters the *Q*-switching regime, and the RF spectrum exhibits again a peak at 17.6 MHz. This behavior is associated with a Hopf bifurcation [41], and consequently, the SL when biased below 218 mA acts as a resonator [21]. The Hopf bifurcation is common in two-section lasers when the SA saturates more easily and at the same time recovers much faster than the gain section [42,43], which in our case can be explained by the small SA-to-gain ratio and the ultrafast dynamics of the QD SA [44]. For  $V_{\text{rev}} = 5.5, 6$  V and low injection current, similar qualitative and quantitative results are observed. The small standard

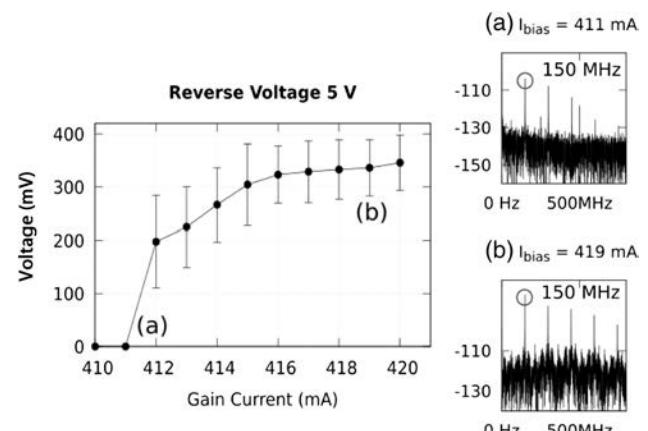
deviation of the spike amplitude, along with the existence of subthreshold oscillations that dictate the spiking frequency, leads to the conclusion that the transition from resting to spiking dynamics is due to a canard explosion, and the system presents a quasi-threshold. This situation has been shown numerically for two-section schemes in Ref. [45].

### B. Neurocomputational Properties Close to the High (Second) GS Current Threshold

As stated above, for  $V_{\text{rev}} \geq 4.5$  V, the SL reveals an additional high-injection GS threshold, along with stable ES emission. For  $V_{\text{rev}} = 5$  V and  $I_{\text{bias}} \approx 410$  mA, the evolution of spike amplitude with the current is shown in Fig. 7. The mean amplitude is close to the case of the low GS threshold, but the standard deviation has increased significantly. For instance, at 412 mA, spike events with amplitude from 0.12 to 0.28 V have been observed, indicating the loss of all-or-nothing dynamics. The RF spectrum at 411 mA reveals a peak at 150 MHz, whose power increases with the gain current (Fig. 7). The dynamical transition is again via a Hopf bifurcation, and the SL below 411 mA acts as a resonator. The same qualitative and quantitative behavior is observed for  $V_{\text{rev}} = 4.5, 5.5, 6$  V, and the proper injection current so as to bias the laser at the “second” GS threshold.

The loss of all-or-nothing dynamics can be linked to the increase of ES emission for higher gain current injection. In particular, the increase of ES photogenerated carriers due to the increased ES emission in the SA can lead to the relaxation of ES carriers to the GS level (the photopumping effect) [46]. Consequently, the GS level of the SA saturates more easily for increased ES emission, and the transition of the peak amplitude with gain current is less abrupt [45]. This slower transition in Hopf bifurcation is associated in neurons with the broadening of the neural threshold area, and consequently, the absence of all-or-nothing dynamics [21].

It is important to notice the increase of the firing rate by 1 order of magnitude (150 MHz compared to 17.6 MHz) in the case of the second GS threshold. Based on the fact that GS-ES interplay can lead to strong *Q*-switching pulsations at multigigahertz frequencies [47] at the onset of dual GS-ES lasing, even for long laser devices, this second GS threshold



**Fig. 7.** Mean amplitude and standard deviation of the GS spike events for biasing close to the high GS threshold and  $V_{\text{rev}} = 5$  V. The marks (a) and (b) correspond to  $I_{\text{bias}} = 411$  and 419 mA.

can potentially unlock neurons with an increased firing rate. This can result in enhanced processing speed and at the same time, adequate optical power for neural network scaling.

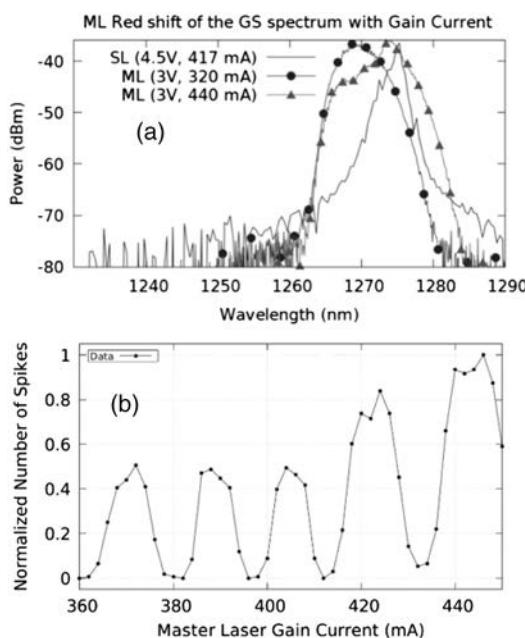
## 5. RESONATE AND FIRE NEURAL OPERATION USING OPTICAL INJECTION FROM A MASTER QD LASER

### A. Optical Excitation with Mode-Locked Pulses

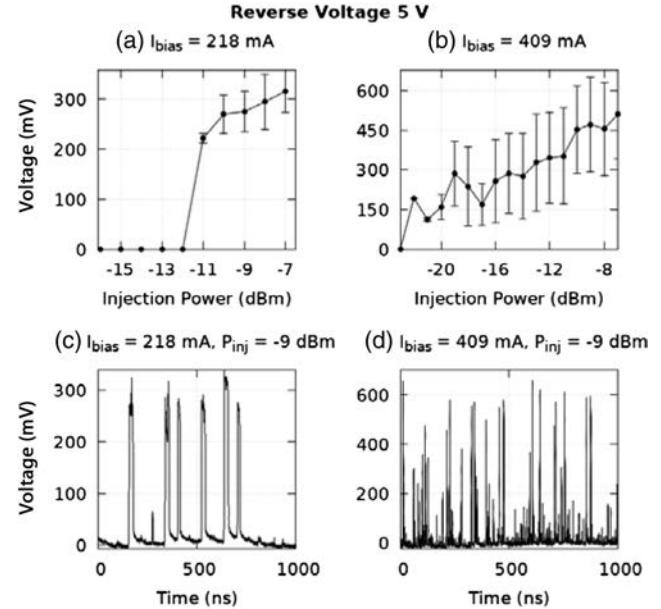
The above analysis revealed the bias conditions under which the QD neuron can act as an integrator or resonator. Nonetheless, an electro-optical-driven laser neuron is of little interest, since the generated neural network is again subject to all the limitations of RF circuits. Therefore, it is vital to provide evidence of all-optical excitability. To this end, the setup is equipped with an ML that provides optical stimuli in the form of picosecond duration mode-locked GS pulses. The ML is biased at 3 V reverse voltage, whereas the gain current can affect simultaneously two experimental parameters: the optical power of the ML pulses and the frequency detuning between the ML and SL. The change of the frequency detuning is caused by the redshift of the ML GS spectrum with increased gain current, as is illustrated in Fig. 8(a) for 320 and 440 mA biasing of the ML. In the same figure, the ML spectrum is compared with the GS spectrum from the SL when it is biased at 4.5 V and 417 mA. The range where the ML current was varied (360–450 mA) did not result in a noticeable power variation, probably due to gain saturation, whereas minor variations were compensated through a polarization controller. It is clear that the low free spectral range of the lasers compared to the resolution of the OSA ( $\Delta\lambda = 0.05$  nm) and the asymmetric, wideband emission

due to the inhomogeneous broadening hinder the definition of a typical ML–SL detuning. Based on this, in Fig. 8(b), it is shown that the optical detuning between the SL–ML's longitudinal modes caused by varying ML current determines the efficiency of the optical injection; this effect is traced through computing the number of spikes triggered at the SL. The SL in this case is biased at 4.5 V. The normalized number of spikes in Fig. 8(b) is calculated by counting the number of spikes in a predefined time window of 50  $\mu$ s duration for all cases. Then the results are normalized with respect to the maximum number of spikes that are observed close to 450 mA. It must be noted that the free spectral range of our Fabry–Perot devices has a resolution below the resolution of the OSA, while the inhomogeneous broadening of the QDs provides a highly asymmetric wideband spectrum, exhibiting multiple peaks [see Fig. 8(a)]. For these reasons, the exact detuning between the ML and the SL could not be computed, and thus the optical injection efficiency is illustrated in terms of the ML gain current in Fig. 8(b).

In Fig. 9, the mean amplitude of spikes as a function of the incoming optical strength is presented when the SL is biased at  $V_{rev} = 5$  V and either at  $I_{bias} = 218$  mA (first GS threshold) or  $I_{bias} = 409$  mA (second GS threshold). Contrary to Fig. 8, where the SL is biased at 4.5 V, in the case of 5 V, the ML achieves the optimum triggering efficiency for ML gain current equal to 400 mA, both in the case where the SL is biased close to its low and high GS thresholds. For 218 mA, the optical injection threshold is observed at -12 dBm, whereas for 409 mA it is observed at -23 dBm. Comparing these two figures, the loss of all-or-nothing threshold is once again observed in the case of the high current biasing.



**Fig. 8.** (a) Detuning between the SL spectrum (black solid line) and the ML spectrum changes with increasing ML gain current (320 mA for the blue line with circles and 480 mA for the red line with triangles); (b) dependence of the number of generated spike events on the ML gain current. The SL is biased at  $V_{rev} = 4.5$  V and  $I_{bias} = 406$  mA, whereas the ML is biased at  $V_{rev} = 3$  V.



**Fig. 9.** Mean amplitude and standard deviation of GS spike events as a function of the injection strength for optical triggering, with the ML biased at 3 V reverse voltage. SL is biased at 5 V reverse voltage and (a) 218 mA—close to the low gain current threshold; (b) 409 mA—close to the high gain current threshold. Time traces for mean injection power equal to -9 dBm are shown for (c) 218 mA and (d) 409 mA.

In Figs. 9(c) and 9(d), time traces are illustrated corresponding to injection strength equal to  $-9$  dBm. For  $I_{\text{bias}} = 218$  mA [Fig. 9(c)], the spike events are all-or-nothing, whereas for  $I_{\text{bias}} = 409$  mA [Fig. 9(d)] it can be seen that all-or-nothing excitability is lost. Moreover, the number of generated spike events is lower in the case of 218 mA as compared to the case of 409 mA, due to the higher firing rate in the second case, as was shown by the RF diagrams.

## B. Multiwaveband Assisted Enhancement of the Temporal Resolution

Another key observation is the variation of the spike's duration versus the existence of ES emission and the reverse bias [48]. In Fig. 10(a), the SL is biased below its low GS current threshold and is triggered by the ML for various reverse bias conditions. At 4.5 V, mean full width at half-maximum (FWHM) is equal to 604 ns, corresponding to optothermal pulsations. Increasing the reverse voltage to 5 V causes the transition from optothermal excitation to  $Q$ -switching-based excitation, which results in a reduction of the mean FWHM by 2 orders of magnitude, to 19 ns. This change can be attributed to the strong

interplay between GS and ES wavebands, as is discussed in Section 4.A. Each subsequent 0.5 V decrease results in a reduction of the mean FWHM by 1 order of magnitude, namely, 2.5 ns for 5.5 V and 0.86 ns for 6 V. The steep reduction with increased voltage may be related not only to the faster carrier escape dynamics in the SA [44], but also to the increased ES emission due to higher unsaturated losses [31].

To shed light on the contribution of ES emission, we reproduce the experiment, but this time, the SL is biased close to the high GS gain threshold, where ES emission is increased [Fig. 10(b)]. The mean FWHM is in the nanosecond scale, even for  $V_{\text{rev}} = 4.5$  V, and weakly varies for increasing reverse voltage. This outcome indicates the significant contribution of ES emission to the temporal resolution. In Fig. 10(c), a 500 ps spike event is illustrated for 5.5 V biasing, which is remarkably close to the 200 ps temporal width provided by the two-section QW-based microlaser [6]. Since the reduction of the cavity size is linked to shorter  $Q$ -switched pulses [49] and a lower ES threshold [31], we expect that the further reduction of cavity size will lead to important improvements regarding the temporal resolution.

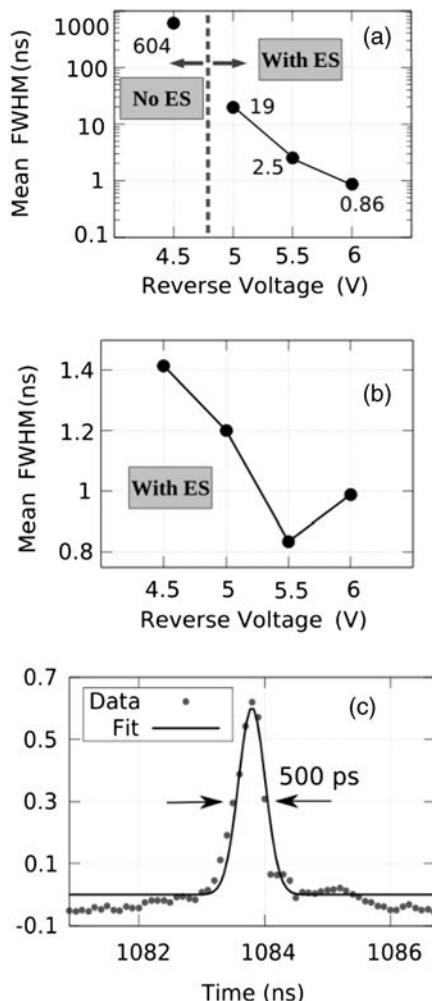
## 6. DISCUSSION

The aforementioned results showcase that the two-section QD laser is able to accomplish both integrate and fire and resonate and fire dynamics, whereas its performance can be enhanced by exploiting the GS-ES interplay. However, the recorded results render the 2 mm long edge-emitting laser adequate only as a proof-of-principle, due to its high energy consumption and the low firing rate achieved compared to other laser-neuron approaches [1], which accomplish a threshold current condition of a few milliamperes and firing rates in the multigigahertz regime. The striking difference between our device and the laser-neurons in the literature can mainly be attributed to the significantly smaller cavity size of the latter (VCSEL devices [7], micropillars [6]).

The main goal of this work is to provide a proof of concept about the new neurocomputational regimes of QD lasers and pave the way for the exploration of these devices as artificial neurons. This fact is enhanced by the fact that the multiwaveband dynamics that are investigated in this work can be transferred to smaller-scale devices, where, in fact, they will be enhanced due to the increased unsaturated losses [32]. Thus, it is anticipated that the design of small-cavity QD lasers (i.e., VCSELs) will provide superior performance as cognitive processors by achieving an extremely low lasing threshold condition due to their 3D confinement structure [50], as well as multigigahertz repetition frequencies [51]. These small-scale devices will still benefit from the thermal stability and the improvements associated with the multiwaveband dynamics that were presented in this and our previous work [20]. These QD-based devices will be able to operate as truly energy-efficient, versatile, and silicon-compatible neuromorphic nodes.

## 7. CONCLUSION

QD lasers, although rarely investigated as potential photonic neural components, have begun attracting notable attention recently due to their thermal insensitivity and their capability of offering seamless integration with silicon photonic platforms.



**Fig. 10.** Spike width as a function of the reverse bias and ES emission (a) close to the low GS current threshold; (b) close to the high GS current threshold. (c) 500 ps spike event recorded for SL with  $I_{\text{bias}} = 407$  mA and  $V_{\text{rev}} = 5.5$  V.

Within this scope, in this work we performed an experimental investigation of the excitable regimes of a multisection QD laser, using injection current in a solitary laser, and optical injection in a unidirectional ML–SL configuration. The multi-section devices are preferred due to the reduced complexity compared to other optical schemes for implementing large-scale neural networks. Results confirm that QD semiconductor lasers can be biased to operate as integrators exploiting the thermo-optic coupling effect, while for higher reverse voltage and in the presence of multiband emission (ES lasing), lasers are driven in a resonate and fire regime. Moreover, a new multi-GS threshold is observed for the first time, originating from the multiwaveband dynamics (GS quenching, ES splitting). The existence of two discrete GS current thresholds is used to demonstrate the significant contribution of the ES emission to the optimization of the temporal resolution. We expect that an increase of ES emission associated with smaller QD devices will further boost these improvements. Consequently, these results, combined with our previous work [20], where the QD laser was studied in terms of its inhibitory and excitatory dynamics, show that the multi-section QD lasers constitute highly flexible devices able to unlock a rich palette of neural attributes such as resonate and fire dynamics, integrate and fire dynamics, along with all-optical excitatory and inhibitory triggering. These properties, which are explicitly related to QDs, when transferred in small-scale devices that benefit from low power consumption, higher temporal resolution, and multigigahertz repetition frequency, will pave the way to the development of truly efficient all-optical spiking neural networks.

**Funding.** Hellenic Foundation for Research and Innovation; General Secretariat for Research and Technology (2247 NEBULA Project).

**Acknowledgment.** The authors would like to thank the members of the photonic technology laboratory in the Department of Informatics and Telecommunications at the University of Athens for hosting the experimental activities of the NEBULA project.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- P. R. Prucnal and B. J. Shastri, *Neuromorphic Photonics* (CRC Press, 2017).
- S. Barland, O. Piro, M. Giudici, J. R. Tredicce, and S. Balle, “Experimental evidence of van der Pol–Fitzhugh–Nagumo dynamics in semiconductor optical amplifiers,” *Phys. Rev. E* **68**, 036209 (2003).
- M. Turconi, B. Garbin, M. Feyereisen, M. Giudici, and S. Barland, “Control of excitable pulses in an injection-locked semiconductor laser,” *Phys. Rev. E* **88**, 022923 (2013).
- B. Garbin, D. Goulding, S. P. Hegarty, G. Huyet, B. Kelleher, and S. Barland, “Incoherent optical triggering of excitable pulses in an injection-locked semiconductor laser,” *Opt. Lett.* **39**, 1254–1257 (2014).
- H. J. Wünsche, O. Brox, M. Radziunas, and F. Henneberger, “Excitability of a semiconductor laser by a two-mode homoclinic bifurcation,” *Phys. Rev. Lett.* **88**, 023901 (2001).
- F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, “Relative refractory period in an excitable semiconductor laser,” *Phys. Rev. Lett.* **112**, 183902 (2014).
- M. A. Nahmias, B. J. Shastri, A. N. Tait, and P. R. Prucnal, “A leaky integrate-and-fire laser neuron for ultrafast cognitive computing,” *IEEE J. Sel. Top. Quantum Electron.* **19**, 1800212 (2013).
- F. Selmi, R. Braive, G. Beaudoin, I. Sagnes, R. Kuszelewicz, and S. Barbay, “Temporal summation in a neuromimetic micropillar laser,” *Opt. Lett.* **40**, 5690–5693 (2015).
- K. Alexander, T. Van Vaerenbergh, M. Fiers, P. Mechet, J. Dambre, and P. Bienstman, “Excitability in optically injected microdisk lasers with phase controlled excitatory and inhibitory response,” *Opt. Express* **21**, 26182–26191 (2013).
- A. M. Yacomotti, P. Monnier, F. Raineri, B. B. Bakir, C. Seassal, R. Raj, and J. A. Levenson, “Fast thermo-optical excitability in a two-dimensional photonic crystal,” *Phys. Rev. Lett.* **97**, 143904 (2006).
- A. Hurtado and J. Javaloyes, “Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems,” *Appl. Phys. Lett.* **107**, 241103 (2015).
- W. W. Chow and F. Jahnke, “On the physics of semiconductor quantum dots for applications in lasers and quantum optics,” *Prog. Quantum Electron.* **37**, 109–184 (2013).
- S. S. Mikhrin, A. R. Kovsh, I. L. Krestnikov, A. V. Kozhukhov, D. A. Livshits, N. N. Ledentsov, Y. M. Shernyakov, I. I. Novikov, M. V. Maximov, and V. M. Ustinov, “High power temperature-insensitive 1.3 μm InAs/InGaAs/GaAs quantum dot lasers,” *Semicond. Sci. Technol.* **20**, 340–342 (2005).
- J. K. Mee, M. T. Crowley, N. Patel, D. Murrell, R. Raghunathan, A. Aboketaf, A. Elshaari, S. F. Preble, P. Ampadu, and L. F. Lester, “A passively mode-locked quantum-dot laser operating over a broad temperature range,” *Appl. Phys. Lett.* **101**, 071112 (2012).
- S. Chen, W. Li, J. Wu, Q. Jiang, M. Tang, S. Shutts, S. N. Elliott, A. Sobiesierski, A. J. Seeds, and I. Ross, “Electrically pumped continuous-wave III–V quantum dot lasers on silicon,” *Nat. Photonics* **10**, 307–311 (2016).
- Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Laroche, and D. Englund, “Deep learning with coherent nanophotonic circuits,” *Nat. Photonics* **11**, 441–446 (2017).
- D. Goulding, S. P. Hegarty, O. Rasskazov, S. Melnik, M. Hartnett, G. Greene, J. G. McInerney, D. Rachinskii, and G. Huyet, “Excitability in a quantum dot semiconductor laser with optical injection,” *Phys. Rev. Lett.* **98**, 153903 (2007).
- M. Dillane, J. Robertson, M. Peters, A. Hurtado, and B. Kelleher, “Neuromorphic dynamics with optically injected quantum dot lasers,” *Eur. Phys. J. B* **92**, 197 (2019).
- G. Sarantoglou, M. Skontranis, and C. Mesaritakis, “All optical integrate and fire neuromorphic node based on single section quantum dot laser,” *IEEE J. Sel. Top. Quantum Electron.* **26**, 1900310 (2020).
- C. Mesaritakis, A. Kapsalis, A. Bogris, and D. Syvridis, “Artificial neuron based on integrated semiconductor quantum dot mode-locked lasers,” *Sci. Rep.* **6**, 39317 (2016).
- E. M. Izhikevich, *Dynamical Systems in Neuroscience* (MIT, 2007).
- R. FitzHugh, “Mathematical models of threshold phenomena in the nerve membrane,” *Bull. Math. Biophys.* **17**, 257–278 (1955).
- M. Krupa and P. Szomolyan, “Relaxation oscillation and canard explosion,” *J. Differ. Equ.* **174**, 312–368 (2001).
- F. Marino, G. Catalán, P. Sánchez, S. Balle, and O. Piro, “Thermo-optical ‘canard orbits’ and excitable limit cycles,” *Phys. Rev. Lett.* **92**, 073901 (2004).
- E. M. Izhikevich, “Resonate-and-fire neurons,” *Neural Netw.* **14**, 883–894 (2001).
- A. Dolcemascolo, B. Garbin, B. Peyce, R. Veltz, and S. Barland, “Resonator neuron and triggering multipulse excitability in laser with injected signal,” *Phys. Rev. E* **98**, 062211 (2018).
- C. Mesaritakis, M. Skontranis, G. Sarantoglou, and A. Bogris, “Micro-ring-resonator based passive photonic spike-time-dependent-plasticity scheme for unsupervised learning in optical neural networks,” in *Optical Fiber Communications Conference and Exhibition (OFC)* (IEEE, 2020), pp. 1–3.
- S. Thorpe, A. Delorme, and R. Van Rullen, “Spike-based strategies for rapid processing,” *Neural Netw.* **14**, 715–725 (2001).

29. J. L. Dubbeldam and B. Krauskopf, "Self-pulsations of lasers with saturable absorber: dynamics and bifurcations," *Opt. Commun.* **159**, 325–338 (1999).
30. J. L. Dubbeldam, B. Krauskopf, and D. Lenstra, "Excitability and coherence resonance in lasers with saturable absorber," *Phys. Rev. E* **60**, 6580–6588 (1999).
31. A. Röhlm, *Dynamic Scenarios in Two-State Quantum Dot Lasers: Excited State Lasing, Ground State Quenching, and Dual-Mode Operation* (Springer, 2015).
32. M. Ishida, N. Hatori, T. Akiyama, K. Otsubo, Y. Nakata, H. Ebe, M. Sugawara, and Y. Arakawa, "Photon lifetime dependence of modulation efficiency and  $K$  factor in 1.3  $\mu\text{m}$  self-assembled InAs/GaAs quantum-dot lasers: impact of capture time and maximum modal gain on modulation bandwidth," *Appl. Phys. Lett.* **85**, 4145–4147 (2004).
33. C. Mesaritakis, C. Simos, H. Simos, A. Kapsalis, E. Roditi, I. Krestnikov, and D. Syvridis, "Effect of the number of quantum dot layers and dual state emission on the performance of InAs/InGaAs passively mode-locked lasers," *Appl. Phys. Lett.* **101**, 251115 (2012).
34. M. Gioannini, "Ground-state power quenching in two-state lasing quantum dot lasers," *J. Appl. Phys.* **111**, 043108 (2012).
35. C. Mesaritakis, C. Simos, H. Simos, I. Krestnikov, and D. Syvridis, "Dual ground-state pulse generation from a passively mode-locked InAs/InGaAs quantum dot laser," *Appl. Phys. Lett.* **99**, 144109 (2011).
36. M. Sugawara, N. Hatori, H. Ebe, M. Ishida, Y. Arakawa, T. Akiyama, K. Otsubo, and Y. Nakata, "Modeling room-temperature lasing spectra of 1.3- $\mu\text{m}$  self-assembled InAs/GaAs quantum-dot lasers: homogeneous broadening of optical gain under current injection," *J. Appl. Phys.* **97**, 043523 (2005).
37. A. Tierro, N. Radwell, and T. Ackemann, "Low-frequency self-pulsing in single-section quantum-dot laser diodes and its relation to optothermal pulsations," *Phys. Rev. A* **84**, 043828 (2011).
38. E. A. Viktorov, T. Erneux, B. Tykalewicz, D. Goulding, S. P. Hegarty, G. Huyet, and B. Kelleher, "Optothermal excitabilities and instabilities in quantum dot lasers," *Proc. SPIE* **9357**, 935704 (2015).
39. E. A. Viktorov and T. Erneux, "Self-sustained pulsations in a quantum-dot laser," *Phys. Rev. E* **90**, 052914 (2014).
40. E. A. Viktorov, M. A. Cataluna, L. O'Faolain, T. F. Krauss, W. Sibbett, E. U. Rafailov, and P. Mandel, "Dynamics of a two-state quantum dot laser with saturable absorber," *Appl. Phys. Lett.* **90**, 121113 (2007).
41. S. H. Strogatz, *Nonlinear Dynamics and Chaos with Student Solutions Manual: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, 2018).
42. T. Erneux and P. Glorieux, *Laser Dynamics* (Cambridge University, 2010).
43. A. G. Vladimirov, U. Bandelow, G. Fiol, D. Arsenijević, M. Kleinert, D. Bimberg, A. Pimenov, and D. Rachinskii, "Dynamical regimes in a monolithic passively mode-locked quantum dot laser," *J. Opt. Soc. Am. B* **27**, 2102–2109 (2010).
44. D. B. Malins, A. Gomez-Iglesias, S. J. White, W. Sibbett, A. Miller, and E. U. Rafailov, "Ultrafast electroabsorption dynamics in an InAs quantum dot saturable absorber at 1.3  $\mu\text{m}$ ," *Appl. Phys. Lett.* **89**, 171111 (2006).
45. G. Kozyreff and T. Erneux, "Singular Hopf bifurcation to strongly pulsating oscillations in lasers containing a saturable absorber," *Eur. J. Appl. Math.* **14**, 407–420 (2003).
46. S. Breuer, M. Rossetti, L. Drzewietzki, P. Bardella, I. Montrosset, and W. Elsasser, "Joint experimental and theoretical investigations of two-state mode locking in a strongly chirped reverse-biased monolithic quantum dot laser," *IEEE J. Quantum Electron.* **47**, 1320–1329 (2011).
47. C. Mesaritakis, C. Simos, H. Simos, A. Kapsalis, I. Krestnikov, and D. Syvridis, "External optical feedback-induced wavelength selection and Q-switching elimination in an InAs/InGaAs passively mode-locked quantum dot laser," *J. Opt. Soc. Am. B* **29**, 1071–1077 (2012).
48. G. Sarantoglou, M. Skontrinis, A. Bogris, and C. Mesaritakis, "Temporal resolution enhancement in quantum-dot laser neurons due to ground state quenching effects," in *Optical Fiber Communication Conference* (Optical Society of America, 2020), paper M2K-1.
49. G. J. Spühler, R. Paschotta, R. Fluck, B. Braun, M. Moser, G. Zhang, E. Gini, and U. Keller, "Experimentally confirmed design guidelines for passively Q-switched microchip lasers using semiconductor saturable absorbers," *J. Opt. Soc. Am. B* **16**, 376–388 (1999).
50. D. G. Deppe, K. Shavritranuruk, G. Ozgur, H. Chen, and S. Freisem, "Quantum dot laser diode with low threshold and low internal loss," *Electron. Lett.* **45**, 54–56 (2009).
51. O. Qasaimeh, W.-D. Zhou, J. Phillips, S. Krishna, P. Bhattacharya, and M. Dutta, "Bistability and self-pulsation in quantum-dot lasers with intracavity quantum-dot saturable absorbers," *Appl. Phys. Lett.* **74**, 1654–1656 (1999).

# PHOTONICS Research

## Engineering of multiple bound states in the continuum by latent representation of freeform structures

RONGHUI LIN,  ZAHRAH ALNAKHLI, AND XIAOHANG LI\* 

*Advanced Semiconductor Laboratory, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Saudi Arabia*

\*Corresponding author: xiaohang.li@kaust.edu.sa

Received 26 November 2020; revised 24 January 2021; accepted 26 January 2021; posted 29 January 2021 (Doc. ID 415655); published 22 March 2021

We demonstrate a neural network capable of designing on-demand multiple symmetry-protected bound states in the continuum (BICs) in freeform structures with predefined symmetry. The latent representation of the freeform structures allows the tuning of the geometry in a differentiable, continuous way. We show the rich band inversion and accidental degeneracy in these freeform structures by interacting with the latent representation directly. Moreover, a high design accuracy is demonstrated for arbitrary control of multiple BIC frequencies by using a photonic property readout network to interpret the latent representation. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415655>

### 1. INTRODUCTION

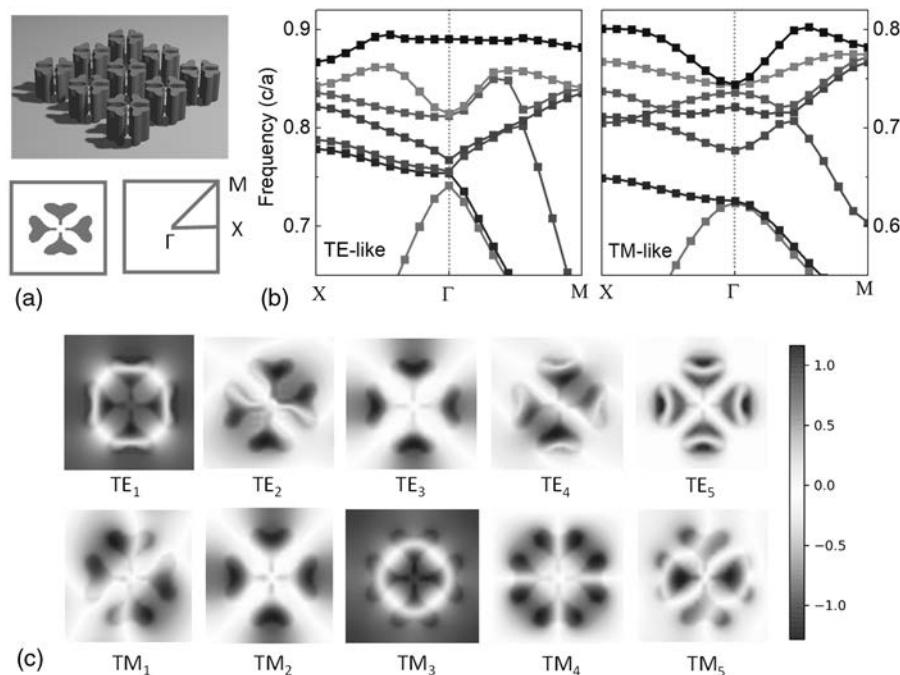
The optical bound states in the continuum (BICs) refer to an exotic class of states that remain perfectly confined despite lying in a continuous spectrum of radiating waves [1,2]. The lack of outgoing radiation means these states have a theoretical infinite lifetime but are unable to be excited by far-field radiation. In practice, perturbations are introduced intentionally or unintentionally to turn the BICs into leaky resonances with finite albeit high  $Q$  factors [3]. Engineering these modes to leverage the high quality ( $Q$ ) factor is benefiting a wide range of applications such as lasers [4,5], sensors [6,7], and nonlinear optics [8,9].

Recent reports show the BICs arise from the vortex centers of the polarization field and they carry quantized topological charges [10]. Merging of multiple BIC points can be used to achieve robust ultra-high  $Q$  factor modes immune to out-of-plane scattering losses [11]. Moreover, being able to manipulate multiple BICs might further benefit areas where high  $Q$  factors are desirable at multiple frequencies such as nonlinear optics [12] and multi-wavelength sensing [13–15].

The core idea of realizing BICs is the parameter tuning to cancel out the far-field radiation. This is mostly achieved by the sweeping of parametric geometries such as circles, rectangles, ellipses, or the combination of them. These geometries can be easily described and modified by equations with a limited number of variables and, thus, a limited degree of tunability. They run into problems with more advanced tuning tasks such as the manipulation of multiple BICs simultaneously. So far, arbitrary control of multiple BICs has never been achieved.

The freeform structures optimized by evolutionary algorithms [16,17] and adjoint methods [18,19] show great

potential in the topology optimization of photonic structures and promise new methods of tuning BICs. They are not bound by any equations and, thus, offer a limitless degree of tunability, which may yield designs that outperform those by conventional geometries [19,20]. However, these algorithms are generally costly in computational resources. Besides these interactive optimization methods, deep neural networks (DNNs) are viable tools in handling complicated photonic structures [20–22]. There have been some impressive attempts to design freeform photonic structures using generative adversarial networks (GANs) [23–27]. A problem with the GANs is the difficulty in training with the possibility of noisy outputs where extra filtering and smoothing algorithms are needed to refine the geometries [24,25]. Another issue with the GANs is that the desired optical properties are directly linked to the geometrical shapes described by pixels, which have much higher dimensions than that of the optical properties such as the transmission and reflection spectra. This results in converging problems and bad generalization performance. Furthermore, the structures with predefined symmetries, which are a critical quality in the field of photonics, have not been demonstrated by GANs. To make the DNNs learn and generate the symmetry and parity properties of the real-life structures is an active research topic in the machine learning community [28–30]. Special techniques such as symmetry loss [31] and structured GANs [32] are necessary to ensure symmetrical outputs. In this work, we demonstrate a DNN structure based on the variational autoencoders (VAEs) that can handle freeform photonic structures with predefined symmetry. Instead of connecting the property to the pixel representation of the geometries directly, we convert the



**Fig. 1.** (a) Top, an artistic rendering of the  $C_{4v}$  photonic crystal considered. Bottom, planar view of the  $C_{4v}$  unit cell and the definition of high symmetry points. (b) Band diagram for TE-like and TM-like modes. (c)  $H_z$  and  $E_z$  Bloch mode profiles for TE-like and TM-like modes, respectively.

geometries into the latent representation, which can be linked to the optical properties more easily and, hence, increase the stability of the inverse design. The latent representation also allows small perturbations of the geometries, which in turn allow continuous manipulation of the photonic properties. With a property readout network to interpret the latent representation, we demonstrate arbitrary, on-demand control of multiple BICs with high accuracy. The band inversion and accidental degeneracy arising from these symmetrical freeform structures can also be a platform for further discoveries and innovations.

## 2. SYMMETRY-PROTECTED BIC IN $C_{4v}$ LATTICE

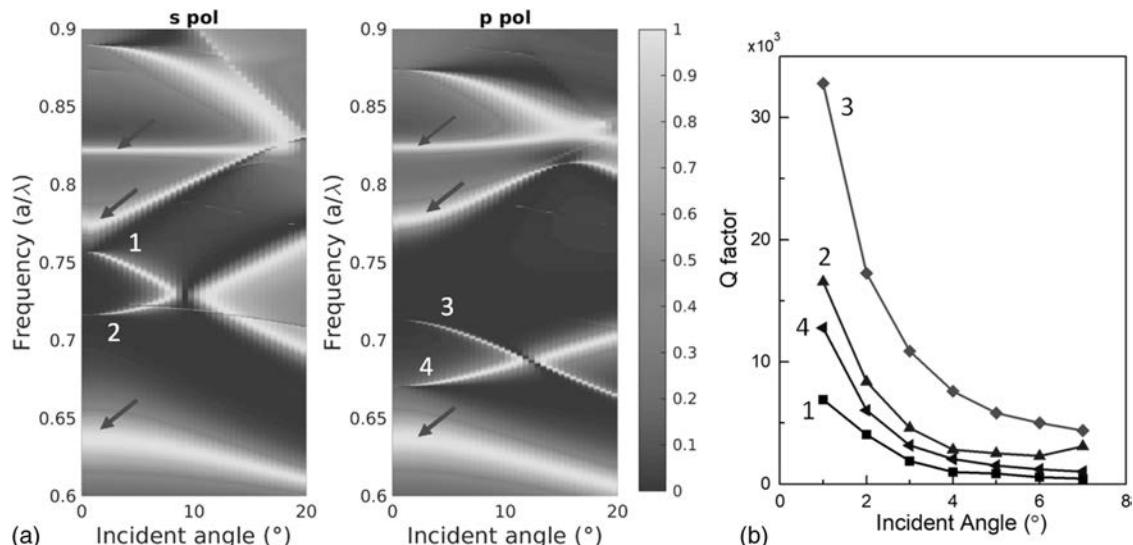
Consider a photonic crystal with  $C_{4v}$  symmetry as shown in Fig. 1(a). At  $\Gamma$  point and below the diffraction limit, the only radiation channel is the zeroth-order diffraction, which has an odd parity under  $C_2$  operation. Therefore, any modes with an even symmetry are decoupled and turned into BICs [1]. As shown in Table 1, the  $C_{4v}$  group has four nondegenerate irreducible representations ( $A_1, A_2, B_1, B_2$ ) and one doubly degenerate irreducible representation ( $E$ ). It is interesting to see that all the nondegenerate representations are even under  $C_2$  transformation. Thus, it is straightforward to conclude that

any nondegenerate modes below the diffraction limit turn into BICs at  $\Gamma$  point.

Since these BICs are protected by symmetry, their frequencies can be tuned by geometrical perturbations as long as the symmetry is maintained. The freeform structure with  $C_{4v}$  symmetry as shown in Fig. 1(a) is studied. We consider a unit cell of  $1 \mu\text{m} \times 1 \mu\text{m} \times 2 \mu\text{m}$  with the material refractive index  $n = 3.5$ , which corresponds to Si in the infrared region. The geometry is assumed to have a fixed height of 550 nm in the air with a mirror symmetry about the  $z$  plane. Hence, the photonic modes can be classified as TE- and TM-like modes. The unit cell is discretized with a  $100 \times 100 \times 200$  mesh, and the plane wave expansion method is used to determine the eigenmodes and eigenfrequencies along the high symmetry points  $X(0, 0.5) - \Gamma(0, 0) - M(0.5, 0.5)$ . The calculations are implemented with an open-source package MIT photonic bands (MPB) [33]. The photonic band diagrams and the corresponding Bloch mode profiles are shown in Figs. 1(b) and 1(c). The  $H_z$  field is plotted for TE-like modes, and the  $E_z$  field is plotted for TM-like modes. The  $TE_2, TE_4, TM_1$ , and  $TM_5$  are doubly degenerate modes, as can be verified from the band diagrams. These mode profiles have an odd parity for  $C_2$  transformation. The nondegenerate modes are all even modes for  $C_2$  transformation, and they all lie below the diffraction limit  $c/na$ , where  $c$ ,  $n$ , and  $a$  are the speed of light, the refractive index of air, and the lattice size, respectively. Hence, they are all symmetry-protected BIC modes. The Bloch modes of the freeform structures largely resemble the field profiles of Mie-type resonators. However, they are distorted and deformed with a lot of local features, which are the key to engineer and fine-tune the BIC frequencies.

**Table 1. Character Table for  $C_{4v}$  Group**

	$E$	$2C_4$	$C_2$	$2\sigma_v$	$2\sigma_d$
$A_1$	1	1	1	1	1
$A_2$	1	1	1	-1	-1
$B_1$	1	-1	1	1	-1
$B_2$	1	-1	1	-1	1
$E$	2	0	-2	0	0



**Fig. 2.** (a) Transmission spectra at different incident angles for s polarization and p polarization. The red arrows indicate the doubly degenerate modes, and the numbered modes are nondegenerate. (b) The  $Q$  factors of the BIC modes shown in (a).

Figure 2(a) shows the transmission spectra at different incident angles for the geometry in Fig. 1 obtained by rigorous coupled-wave analysis (RCWA). The  $0^\circ$  incident angle corresponds to the  $\Gamma$  point. We can observe three modes at frequencies 0.63, 0.77, 0.83 in both s polarization and p polarization [marked by red arrows in Fig. 2(a)]. They correspond to the TM<sub>1</sub>, TE<sub>2</sub>, and TE<sub>4</sub> doubly degenerate modes because the s polarization and p polarization have a  $90^\circ$  rotation relationship, which matches the symmetries of the doubly degenerate modes shown in Fig. 1. Other modes, such as the numbered ones in Fig. 2(a), are nondegenerate, which appear either in s polarization or p polarization. These modes show vanishing linewidths as the incident angle approaches  $0^\circ$ . The  $Q$  factors for the four modes highlighted show diverging trends as the incident angle approaches  $0^\circ$ , which indicates that they are indeed BICs.

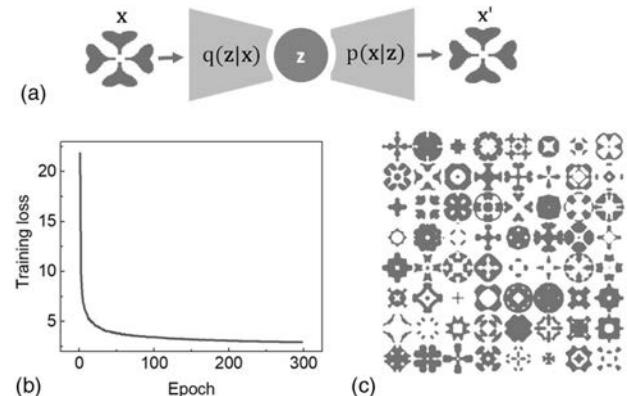
### 3. LATENT REPRESENTATION OF GEOMETRY USING $\beta$ -VAE

A VAE structure as shown in Fig. 3 is designed to manage the freeform structures with  $C_{4v}$  symmetry. In the VAE, the input  $\mathbf{x}$  passes through two DNNs sequentially, namely the encoder  $q(\mathbf{z}|\mathbf{x})$  and the decoder  $p(\mathbf{x}|\mathbf{z})$ , to output a probability distribution  $\mathbf{x}'$ . The latent vector  $\mathbf{z}$  has a much lower dimension than the input  $\mathbf{x}$ . As the data pass through such a bottleneck layer, they are compressed, and the network is forced to learn an efficient representation of the input. Data in a much higher dimension can be represented by a low dimension vector faithfully after training. In this study, the input is  $64 \times 64$  pixels binary images with 1 representing the material and 0 representing the air. The latent vector  $\mathbf{z}$  is an array with the form  $[z_1, z_2, \dots, z_{10}]$ .

We use the  $\beta$ -VAE [34,35] with a loss function as follows:

$$\mathcal{L} = -\mathbb{E}_{q(\mathbf{z}|\mathbf{x})}[\log p(\mathbf{x}|\mathbf{z})] + \beta D_{KL}[q(\mathbf{z}|\mathbf{x})||p(\mathbf{z})]. \quad (1)$$

The first term is the reconstruction loss, which forces the decoder to represent the input as closely as possible. The second



**Fig. 3.** (a) VAE structure used for geometry management. (b) The training loss for  $\beta$ -VAE. (c) Examples of the  $\beta$ -VAE generated geometries.

term is the Kullback–Leibler (KL) divergence between the prior distribution  $p(\mathbf{z})$  and the encoder distribution  $q(\mathbf{z}|\mathbf{x})$ . It is a regularization term that forces the latent representation to assume the same standard normal distribution as the prior  $p(\mathbf{z})$ . The resultant latent representation is centered and closely packed in the latent space, and most importantly, continuous and interpolable. The  $\beta$  factor imposes extra weight on the KL divergence, thus increasing the regularization power. Previous reports show an increase in the  $\beta$  factor can promote the disentanglement of latent representation [35], allowing them to be more interpretable. In this work, we choose a  $\beta$  factor of 3. The  $\beta$ -VAE can not only reproduce the training data but also create new data with the same distribution as the training data. As the high dimension data are collapsed to a lower dimension, a lot of the high dimension features and minor details are filtered out. Hence, the output geometries are naturally smooth with less ultrafine structures such as sharp corners and isolated islands. Moreover, the latent representation has similar dimensions with

our optical response, which is easier to link to the optical properties by DNN.

We use 20,000 randomly generated geometries with  $C_{4v}$  symmetry to train the  $\beta$ -VAE. The training geometries are generated by applying symmetry operations to random polygons. The training loss of the  $\beta$ -VAE in Fig. 3(b) shows good convergence, which is an advantage over GANs. Randomly generated latent vectors are used to test the output of the  $\beta$ -VAE after the training. Some examples of the output are shown in Fig. 3(c). The VAE can produce a wide variety of geometries with different topology while maintaining the perfect  $C_{4v}$  symmetry. These geometries are obtained without any additional filtering, but they are smooth with little noise. Since the latent representation is continuous, a perturbation of the latent vector  $\Delta\mathbf{z}$  results in a perturbation  $\Delta\mathbf{x}'$  in the output. This is demonstrated in Visualization 1, where we show the change of the output geometries as the  $\mathbf{z}$  vector is varied continuously.

To study how the continuous deformation of such freeform structures influences the photonic band structures, we generate a random initial geometry and then vary one component of the latent vector  $z_3$  continuously with a step of 0.02. A total number of 20 geometries are generated, and geometries 8 and 12 are compared in detail. Geometry 8 is the same as the one depicted in Fig. 1, and geometry 12 is a slightly deformed version. Their shapes are shown in Fig. 4(a). The band frequencies at the  $\Gamma$  point shift smoothly when the geometrical shapes change continuously, as shown in Fig. 4(b). An interesting phenomenon is the band crossing during this process. Comparing with the TE Bloch mode profiles in Fig. 1, we find the mode profiles of  $TE_2$  and  $TE_3$  are swapped. Between these two shapes, there is an intermediate shape (geometry 10) where these two bands cross and form an “accidental degeneracy.”  $TE_2$  is a nondegenerate mode with even parity while  $TE_3$  is a doubly degenerate mode with odd parity. Previous studies show that a Dirac cone with linear dispersion [36,37] appears when these bands meet. Such a Dirac cone dispersion plays a vital role in zero-index materials

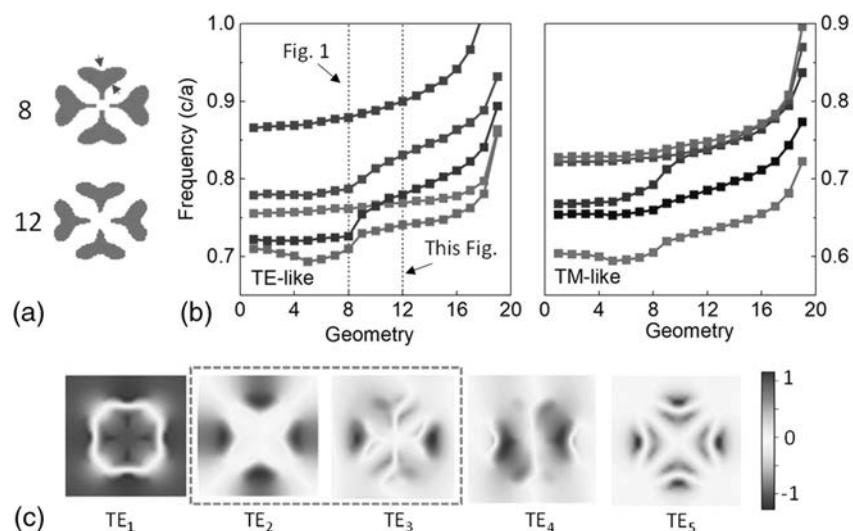
[38,39] and topological photonics [40], which could be another application for these freeform structures.

The reason for this band crossing can be understood from the Bloch mode profiles. In Fig. 4(c), the field of  $TE_2$  is mostly localized at the center of the heart-shaped structures; hence, it is less sensitive as the heart shape deforms. In contrast,  $TE_3$  is more susceptible to such changes because a considerable amount of field intensity is located at the site of deformation. Since the amount of the frequency shift is proportional to the field concentration in the perturbation area [41], the  $TE_2$  band remains mostly unchanged while the  $TE_3$  band shifts more drastically, causing the band crossing. In this geometry,  $TE_1$ ,  $TE_2$ , and  $TE_5$  are symmetry-protected BICs. The shifting and crossing of the bands give us a lot of possible combinations of BICs. Since these shapes are controlled by a latent vector with 10 dimensions, there are 10 dimensions to fine-tune the geometries and, hence, the BIC combinations.

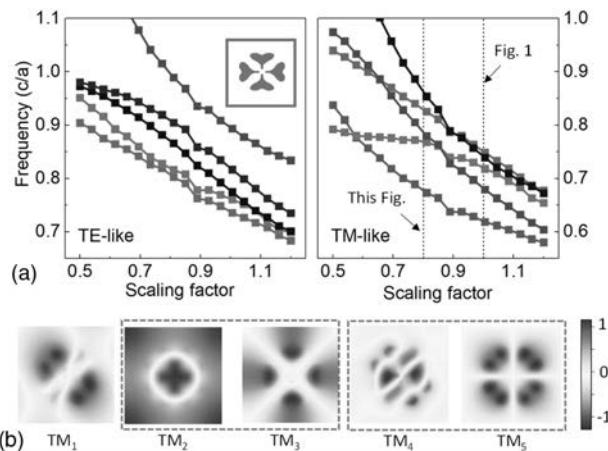
We also study the shift of the photonic bands as the geometry in Fig. 1 is scaled by a factor of 0.5 to 1.2. The band frequencies at  $\Gamma$  point are shown in Fig. 5(a), and the Bloch mode profiles are shown in Fig. 5(b). Comparing the band order to that of Fig. 1, we notice drastic band crossing for TM polarization. Two sets of band crossing [( $TM_2$ ,  $TM_3$ ) and ( $TM_4$ ,  $TM_5$ )] are observed when the scaling factor is varied from 1 to 0.8.  $TM_4$  has odd parity, and  $TM_5$  has even parity while both  $TM_2$  and  $TM_3$  have even parity. The crossing of  $TM_4$  and  $TM_5$  results in a Dirac-like dispersion while the crossing of  $TM_2$  and  $TM_3$  would lead to a double quadratic dispersion as shown in previous reports [36]. These results show great potential for band engineering for photonic applications.

#### 4. ON-DEMAND MULTIPLE BIC DESIGN

The continuous deformation and scaling of the geometry make it possible to engineer a large number of BICs. To achieve the on-demand design of multiple BICs, a property readout

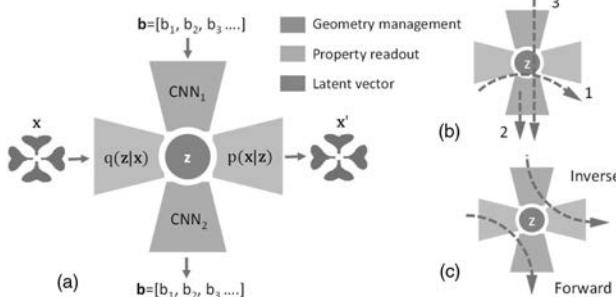


**Fig. 4.** (a) Shapes of geometry 8 and 12, with the sites of deformation marked by red arrows. (b) The shift of TE-like and TM-like bands at  $\Gamma$  point as the latent vector is varied continuously (see Visualization 1 for the continuous variation of the geometries). (c)  $H_z$  field of TE-like modes for geometry 12. The inversed bands are grouped by dashed green boxes.



**Fig. 5.** (a) Shift of TE-like and TM-like bands at  $\Gamma$  point as the scaling factor is varied. The inset in (a) shows the geometry considered. (b) The  $E_z$  field of TM-like modes for geometry with a scaling factor of 0.8. The inverted bands are grouped by dashed green boxes.

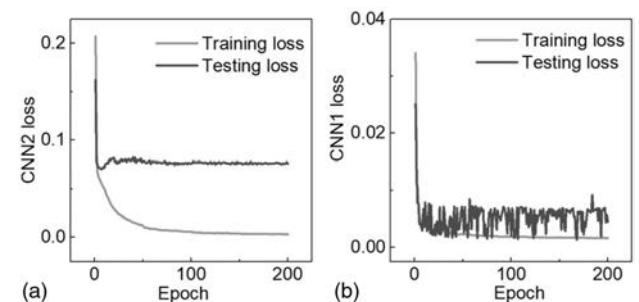
network is necessary to link the latent representation of geometries to their photonic properties. The whole DNN structure is shown in Fig. 6(a). The property readout network, consisting of two convolutional neural networks (CNNs), is connected to the geometry handling network ( $\beta$ -VAE) by the latent vector  $\mathbf{z}$ . The information flow during training is shown in Fig. 6(b). In process 1, the geometry management network is trained to realize the mapping between the input geometry  $\mathbf{x}$ , latent space  $\mathbf{z}$ , and output geometry  $\mathbf{x}'$ . With this mapping established, a total number of 20,000  $\mathbf{z}$  vectors are sampled from the latent space, and their corresponding geometries are recorded. After that, their photonic bands at the  $\Gamma$  point are calculated. Then the BIC frequencies are selected by removing the degenerate modes. Three lowest BIC frequencies are picked to form the BIC vector  $\mathbf{b} = [b_1, b_2, b_3]$ . The  $\mathbf{z}$ - $\mathbf{b}$  pair is used to train the property readout network. In process 2, the latent vector  $\mathbf{z}$  is set as the input and  $\mathbf{b}$  as the output to train CNN<sub>2</sub>. This ensures the forward mapping from the latent vector  $\mathbf{z}$  to its corresponding  $\mathbf{b}$  vector. After that, the weights of CNN<sub>2</sub> are fixed when the BIC vector  $\mathbf{b}$  is fed to the CNN<sub>1</sub>- $\mathbf{z}$ -CNN<sub>2</sub> network, and the weights of CNN<sub>1</sub> are updated. This CNN<sub>1</sub>- $\mathbf{z}$ -CNN<sub>2</sub> structure is similar to the tandem network [42] that is used to solve



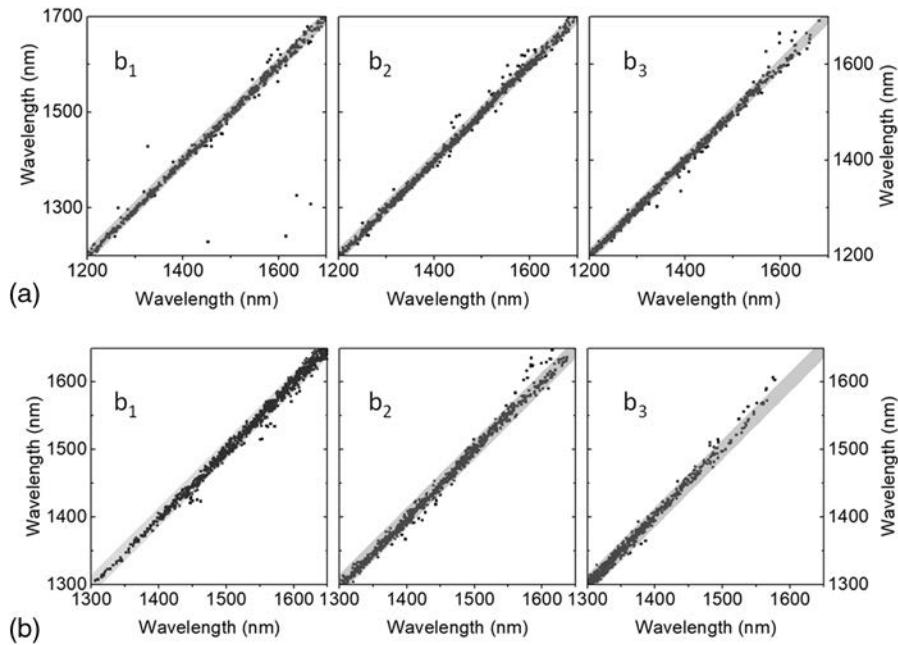
**Fig. 6.** (a) Whole DNN structure studied in this work. (b) Information flow during training, where 1 is the training of  $\beta$ -VAE, 2 is the training of CNN<sub>2</sub>, and 3 is the training of CNN<sub>1</sub>. (c) Information flow for the forward modeling and the inverse design.

data inconsistency problems. Since the mapping from physical property  $\mathbf{b}$  to the latent vector  $\mathbf{z}$  is not unique, there might be converging problems if CNN<sub>1</sub> is trained directly. The idea is that the mapping from  $\mathbf{z}$  to  $\mathbf{b}$  is always one to one so that CNN<sub>2</sub> can always converge to one possible solution. After that, this solution is used as the criteria to train CNN<sub>1</sub>, which ensures the convergence. After all the networks are trained, the whole DNN can be used for the forward modeling or the inverse design, as shown in Fig. 6(c). In the former case, a geometry is fed to the network, and the output is the BIC vector  $\mathbf{b}$ . In the latter case, a target  $\mathbf{b}$  vector is fed to the network, and the output is the geometry. The detailed parameters of the whole DNN can be found in the last section of this work. We use 18,000 sets of  $\mathbf{z}$ - $\mathbf{b}$  pairs for training, 1000 for testing, and the final 1000 for validation. The mean squared error (MSE) is used for the training of both CNN<sub>1</sub> and CNN<sub>2</sub>. The training and testing losses are shown in Fig. 7. It can be seen that the training and testing losses converge after 200 epochs for both CNN<sub>1</sub> and CNN<sub>2</sub>. We use both the validation data set and randomly generated data to test the performance of the network. In Fig. 8(a), the 1000 sets of reserved validation data are fed to the CNN<sub>1</sub>- $\mathbf{z}$ -CNN<sub>2</sub> network, and the output is compared to the target value. In Fig. 8(b), we generate 1000 random  $\mathbf{b}$  vectors to test the output of the network. In both figures, the  $x$  axis is the target value, and the  $y$  axis is the output of the DNN. The green bands are defined by  $y = x \pm 15$  nm. We can see a strong linear correlation for the target value and the network output with 99% of the points lying within the bands of  $y = x \pm 15$  nm, indicating good design accuracy, even for randomly generated input.

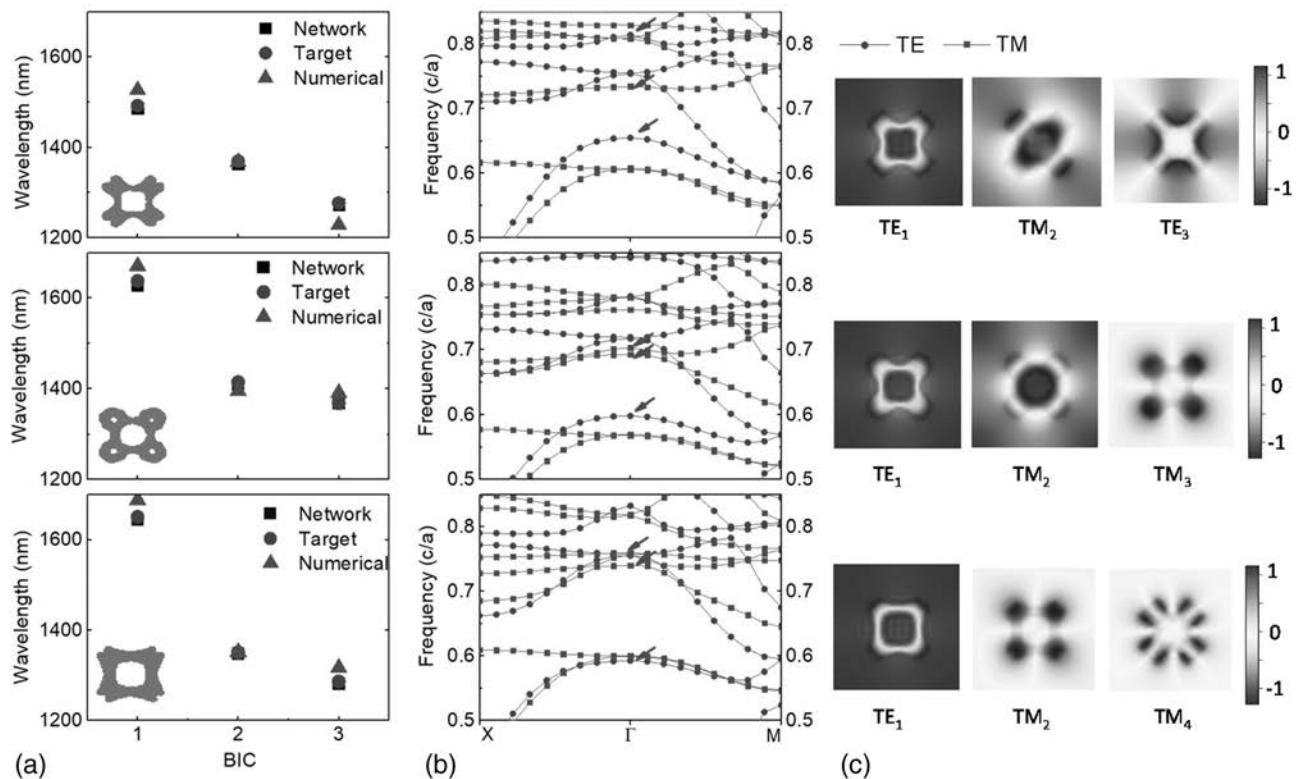
Next, we show a few specific examples of the inverse design of multiple BICs. Three sets of random  $\mathbf{b}$  vectors are generated and fed to the DNN through the inverse design pathway. The obtained geometries are simulated to determine their BIC frequencies. The target value, the value inferred by DNN, and the value obtained by simulations agree well, as shown in Fig. 9(a). The insets show the output geometries. Figure 9(b) shows the band diagrams for the output geometries with the target BIC frequencies marked by red arrows. Figure 9(c) shows the corresponding mode profiles at the BIC frequencies. The BIC frequencies are all nondegenerate modes with an even parity for the  $C_2$  transformation, which means they are indeed BIC states. We observe the maximum discrepancy for  $b_1$  and  $b_3$  is around 30 nm, which is within the  $\pm 15$  nm range shown earlier. This discrepancy is around 2.5% of the shortest operation wavelength (from 1200 nm to 1700 nm), which is



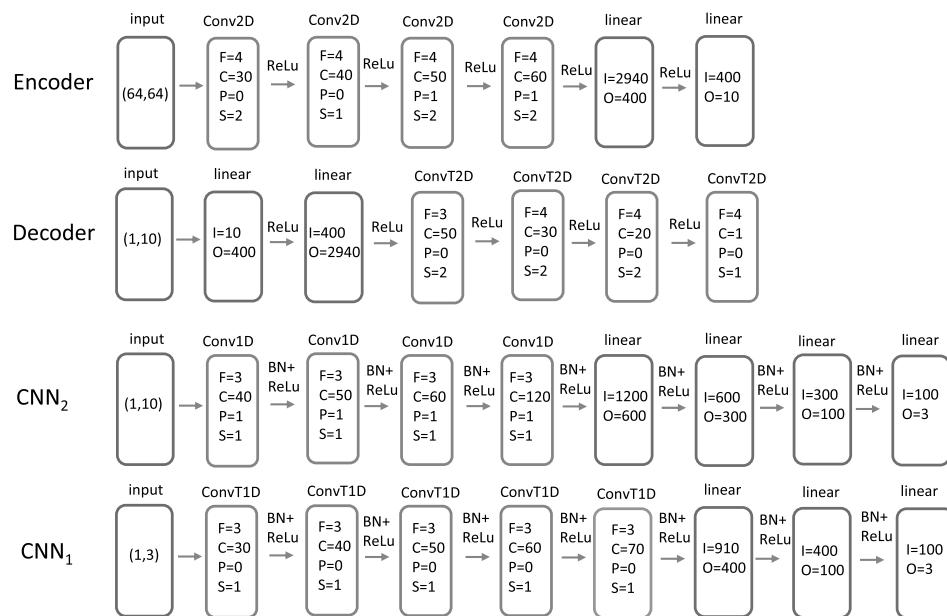
**Fig. 7.** Training and testing losses for (a) CNN<sub>2</sub> and (b) CNN<sub>1</sub>.



**Fig. 8.** (a) Correlation between target BIC wavelengths in the validation data set and the output of the DNN. (b) Correlation between randomly generated target BIC wavelengths and the output of the DNN. The  $x$  axis is the target value, and the  $y$  axis is the DNN output in both (a) and (b).



**Fig. 9.** Demonstration of the multiple BIC inverse design. (a) The comparison of random target BIC wavelengths, DNN output, and numerical simulation. The inset shows the output geometry of the DNN. (b) Band diagrams for the designed structures. (c) The  $H_z$  and  $E_z$  field profiles for the TE and TM BIC states, respectively.



**Fig. 10.** Detailed parameters of the network structure.

adequate in actual applications. Moreover, a higher design accuracy can be obtained by higher pixel resolution in the image representation.

## 5. PARAMETERS OF THE DNN

Figure 10 shows the parameters of the DNN architecture. The encoder contains four convolutional (Conv2D) layers and two linear layers. Its input is a  $64 \times 64$  pixels image with a pixel value of 0 or 1. For convolutional layers,  $F$  is the filter size,  $C$  is the number of channels,  $P$  is the padding size, and  $S$  is the stride.  $I$  and  $O$  are the input and the output for linear layers. ReLu stands for rectified linear unit activation function. The decoder consists of two linear layers and four transposed convolution layers (ConvT2D). Both the decoder and encoder are based on CNN structures developed for 2D image classification [43]. The CNN<sub>2</sub> has an input dimension of (1, 10), and it consists of four convolutional layers (Conv1D) and four linear layers. A batch normalization layer (BN) is applied before the ReLu to increase the convergence and generalization of the network. The CNN<sub>1</sub> has an input dimension of (1, 3), and it consists of five ConvT1D layers and three linear layers. The network is implemented in Python using the Pytorch machine learning library. The parameters of CNN<sub>1</sub> and CNN<sub>2</sub> are based on an early work with similar data structures [22].

## 6. CONCLUSION

We demonstrate a DNN structure that can design and engineer multiple symmetry-protected BICs by manipulating freeform structures with predefined symmetry. The geometries are represented by latent vectors, which can then be mapped to the photonic property by a property readout network. We demonstrate the on-demand design of three arbitrary BIC frequencies with high accuracy. We also analyze the nature of complicated band inverse and accidental degeneracy when such freeform

structures are tuned and scaled continuously, which shows the potential for further discovery and application.

**Funding.** King Abdullah University of Science and Technology Baseline Fund (BAS/1/1664-01-01); Competitive Research Grants (URF/1/3437-01-01, URF/1/3771-01-01).

**Acknowledgment.** The authors would like to thank Prof. Andrea Fratalocchi from KAUST for his advice.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

1. C. W. Hsu, B. Zhen, A. D. Stone, J. D. Joannopoulos, and M. Soljačić, “Bound states in the continuum,” *Nat. Rev. Mater.* **1**, 16048 (2016).
2. C. W. Hsu, B. Zhen, J. Lee, S.-L. Chua, S. G. Johnson, J. D. Joannopoulos, and M. Soljačić, “Observation of trapped light within the radiation continuum,” *Nature* **499**, 188–191 (2013).
3. K. Koshelev, S. Lepeshov, M. Liu, A. Bogdanov, and Y. Kivshar, “Asymmetric metasurfaces with high-Q resonances governed by bound states in the continuum,” *Phys. Rev. Lett.* **121**, 193903 (2018).
4. A. Kodigala, T. Lepetit, Q. Gu, B. Bahari, Y. Fainman, and B. Kanté, “Lasing action from photonic bound states in continuum,” *Nature* **541**, 196–199 (2017).
5. C. Huang, C. Zhang, S. Xiao, Y. Wang, Y. Fan, Y. Liu, N. Zhang, G. Qu, H. Ji, J. Han, L. Ge, Y. Kivshar, and Q. Song, “Ultrafast control of vortex microlasers,” *Science* **367**, 1018–1021 (2020).
6. M. Meudt, C. Bogiadzi, K. Wrobel, and P. Görrn, “Hybrid photonic-plasmonic bound states in continuum for enhanced light manipulation,” *Adv. Opt. Mater.* **8**, 2000898 (2020).
7. F. Yesilkoy, E. R. Arvelo, Y. Jahani, M. Liu, A. Tittl, V. Cevher, Y. Kivshar, and H. Altug, “Ultrasensitive hyperspectral imaging and bio-detection enabled by dielectric metasurfaces,” *Nat. Photonics* **13**, 390–396 (2019).
8. Z. Liu, Y. Xu, Y. Lin, J. Xiang, T. Feng, Q. Cao, J. Li, S. Lan, and J. Liu, “High-Q quasibound states in the continuum for nonlinear metasurfaces,” *Phys. Rev. Lett.* **123**, 253901 (2019).

9. L. Xu, K. Zangeneh Kamali, L. Huang, M. Rahmani, A. Smirnov, R. Camacho-Morales, Y. Ma, G. Zhang, M. Woolley, D. Neshev, and A. E. Miroshnichenko, "Dynamic nonlinear image tuning through magnetic dipole quasi-BIC ultrathin resonators," *Adv. Sci.* **6**, 1802119 (2019).
10. B. Zhen, C. W. Hsu, L. Lu, A. D. Stone, and M. Soljačić, "Topological nature of optical bound states in the continuum," *Phys. Rev. Lett.* **113**, 257401 (2014).
11. J. Jin, X. Yin, L. Ni, M. Soljačić, B. Zhen, and C. Peng, "Topologically enabled ultrahigh-Q guided resonances robust to out-of-plane scattering," *Nature* **574**, 501–504 (2019).
12. Y. W. Lee, F. C. Fan, Y. C. Huang, B. Y. Gu, B. Z. Dong, and M. H. Chou, "Nonlinear multiwavelength conversion based on an aperiodic optical superlattice in lithium niobate," *Opt. Lett.* **27**, 2191–2193 (2002).
13. S. Diaz, "Stable dual-wavelength erbium fiber ring laser with optical feedback for remote sensing," *J. Lightwave Technol.* **34**, 4591–4595 (2016).
14. Y. Zuo, Y. Gao, S. Qin, Z. Wang, D. Zhou, Z. Li, Y. Yu, M. Shao, and X. Zhang, "Broadband multi-wavelength optical sensing based on photothermal effect of 2D MXene films," *Nanophotonics* **9**, 123–131 (2019).
15. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
16. W.-K. Lee, S. Yu, C. J. Engel, T. Reese, D. Rhee, W. Chen, and T. W. Odom, "Concurrent design of quasi-random photonic nanostructures," *Proc. Natl. Acad. Sci. USA* **114**, 8734–8739 (2017).
17. T. Feichtner, O. Selig, M. Kiunke, and B. Hecht, "Evolutionary optimization of optical antennas," *Phys. Rev. Lett.* **109**, 127701 (2012).
18. Z. Shi, A. Y. Zhu, Z. Li, Y.-W. Huang, W. T. Chen, C.-W. Qiu, and F. Capasso, "Continuous angle-tunable birefringence with freeform metasurfaces for arbitrary polarization conversion," *Sci. Adv.* **6**, eaba3367 (2020).
19. A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nat. Photonics* **9**, 374–377 (2015).
20. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**, 339–366 (2019).
21. R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.* **2**, 1007–1023 (2020).
22. R. Lin, Y. Zhai, C. Xiong, and X. Li, "Inverse design of plasmonic metasurfaces by convolutional neural network," *Opt. Lett.* **45**, 1362–1365 (2020).
23. T. Christensen, C. Loh, S. Picek, D. Jakobović, L. Jing, S. Fisher, V. Ceperic, J. D. Joannopoulos, and M. Soljačić, "Predictive and generative machine learning models for photonic crystals," *Nanophotonics* **9**, 4183–4192 (2020).
24. Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
25. J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Freeform diffractive metagrating design based on generative adversarial networks," *ACS Nano* **13**, 8872–8878 (2019).
26. F. Wen, J. Jiang, and J. A. Fan, "Progressive-growing of generative adversarial networks for metasurface optimization," arXiv:1911.13029 (2019).
27. S. So and J. Rho, "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**, 1255–1261 (2019).
28. T. Cohen and M. Welling, "Group equivariant convolutional networks," in *International Conference on Machine Learning* (2016), pp. 2990–2999.
29. S. Dieleman, J. De Fauw, and K. Kavukcuoglu, "Exploiting cyclic symmetry in convolutional neural networks," arXiv:1602.02660 (2016).
30. I. Nachum and A. Yehudayoff, "On symmetry and initialization for neural networks," arXiv:1907.00560 (2019).
31. R. Huang, S. Zhang, T. Li, and R. He, "Beyond face rotation: global and local perception GAN for photorealistic and identity preserving frontal view synthesis," in *IEEE International Conference on Computer Vision* (2017), pp. 2439–2448.
32. I. Peleg and L. Wolf, "Structured GANs," in *IEEE Winter Conference on Applications of Computer Vision (WACV)* (2018), pp. 719–728.
33. S. G. Johnson and J. D. Joannopoulos, "Block-iterative frequency-domain methods for Maxwell's equations in a planewave basis," *Opt. Express* **8**, 173–190 (2001).
34. I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. M. Botvinick, S. Mohamed, and A. Lerchner, "Beta-VAE: learning basic visual concepts with a constrained variational framework," in *International Conference on Learning Representations* (2017), pp. 1–13.
35. C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner, "Understanding disentangling in β-VAE," arXiv:1804.03599 (2018).
36. K. Sakoda, "Universality of mode symmetries in creating photonic Dirac cones," *J. Opt. Soc. Am. B* **29**, 2770–2778 (2012).
37. Y. Li, Y. Wu, X. Chen, and J. Mei, "Selection rule for Dirac-like points in two-dimensional dielectric photonic crystals," *Opt. Express* **21**, 7699–7711 (2013).
38. X. Huang, Y. Lai, Z. H. Hang, H. Zheng, and C. T. Chan, "Dirac cones induced by accidental degeneracy in photonic crystals and zero-refractive-index materials," *Nat. Mater.* **10**, 582–586 (2011).
39. M. Minkov, I. A. D. Williamson, M. Xiao, and S. Fan, "Zero-index bound states in the continuum," *Phys. Rev. Lett.* **121**, 263901 (2018).
40. L. Lu, J. D. Joannopoulos, and M. Soljačić, "Topological photonics," *Nat. Photonics* **8**, 821–829 (2014).
41. J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University, 2011).
42. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
43. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM* **60**, 84–90 (2017).

# PHOTONICS Research

## Real-time deep learning design tool for far-field radiation profile

JINRAN QIE,<sup>1</sup> ERFAN KHORAM,<sup>2</sup> DIANJING LIU,<sup>2</sup> MING ZHOU,<sup>2</sup> AND LI GAO<sup>3,\*</sup>

<sup>1</sup>Department of Electrical and Systems Engineering, Washington University, St Louis, Missouri 63130, USA

<sup>2</sup>Department of Electrical and Computer Engineering, University of Wisconsin, Madison, Wisconsin 53706, USA

<sup>3</sup>Key Laboratory for Organic Electronics & Information Displays (KLOEID), Institute of Advanced Materials (IAM), and School of Materials Science and Engineering, Nanjing University of Posts & Telecommunications, Nanjing 210046, China

\*Corresponding author: iamlgao@njupt.edu.cn

Received 27 October 2020; revised 28 December 2020; accepted 2 February 2021; posted 2 February 2021 (Doc. ID 413567); published 23 March 2021

The connection between Maxwell's equations and artificial neural networks has revolutionized the capability and efficiency of nanophotonic design. Such a machine learning tool can help designers avoid iterative, time-consuming electromagnetic simulations and even allows long-desired inverse design. However, when we move from conventional design methods to machine-learning-based tools, there is a steep learning curve that is not as user-friendly as commercial simulation software. Here, we introduce a real-time, web-based design tool that uses a trained deep neural network (DNN) for accurate far-field radiation prediction, which shows great potential and convenience for antenna and metasurface designs. We believe our approach provides a user-friendly, readily accessible deep learning design tool, with significantly reduced difficulty and greatly enhanced efficiency. The web-based tool paves the way to present complicated machine learning results in an intuitive way. It also can be extended to other nanophotonic designs based on DNNs and replace conventional full-wave simulations with a much simpler interface. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.413567>

### 1. INTRODUCTION

Nanophotonic devices offer new capabilities to control light with nanostructures designed for different functionalities. In these photonic devices, a large number of geometric parameters play critical roles in altering the light–matter interaction. For complex nanostructures, there could be millions or even billions of combinations of all possible structures. However, conventional design methods rely on time-consuming, full-wave simulations and an iterative optimization process. It is challenging to explore all the options; usually only a limited number of designs are explored, leaving an enormous parameter space underexplored.

Machine learning has led to revolutionary developments in numerous applications. Its complex models and algorithms can help exploit the enormous parameter space in nanophotonics, enabling both efficient forward prediction and on-demand inverse designs. Artificial neural networks (ANNs) [1–6] are an interconnected group of nodes that are similar to the complicated network of neurons in a brain with the capability of self-learning. It is a data-driven approach, which is in contrast to a computation-driven approach, such as optimization [7–10]. Recent representative examples include near- and far-field prediction [11], metasurface and metamaterials designs [12–14],

and structural color design [15]. The merging of deep learning and nanophotonics has reduced computation time by orders of magnitude and expands the design space that previously could not be realized.

In Ref. [11], the authors show good agreements of near-field and far-field scattering of three-dimensional (3D) nanostructures between the simulation results and the neural network prediction by using a convolutional neural network. Unfortunately, the transition from conventional simulation to deep-learning-based tools requires knowledge from both nanophotonics and computation. There is a steep learning curve that hinders researchers from accessing such a convenient and efficient tool. On the other hand, there is no demonstration of the application of deep neural networks (DNNs) for nanophotonic design in a web-based, real-time setting. If a trained DNN can be interfaced with a web-page tool, the real application process can be even simpler than using commercial simulation software, with accurate results displayed in real time. In this way, device designers can effortlessly benefit from deep-learning-enabled computation.

Here, we demonstrate the training of DNNs for accurate far-field pattern prediction of dielectric antennas. Then we interface the DNNs with a web-page tool for real-time design output. The far-field radiation profile [16] is used in the design

of many nanophotonic devices, such as optical antennas [17,18] and metasurfaces [19–21]. The conventional approach to obtain a far-field radiation pattern is through finite difference time domain (FDTD) simulation with near- to far-field transformation [22,23], or using a commercial/open source software package; either method can take from tens of minutes to hours to complete the computation. In practice, iterative optimization is necessary to get an optimal structure. For repetitive designs, it requires a considerable amount of computation resource and time. By investing one set of DNNs training data, we demonstrate an online tool to predict the far-field radiation pattern of any arbitrary scatterer in real time. The results suggest that a web-page tool can maximize the advantages of a DNN-based design method and significantly improve a designer's productivity.

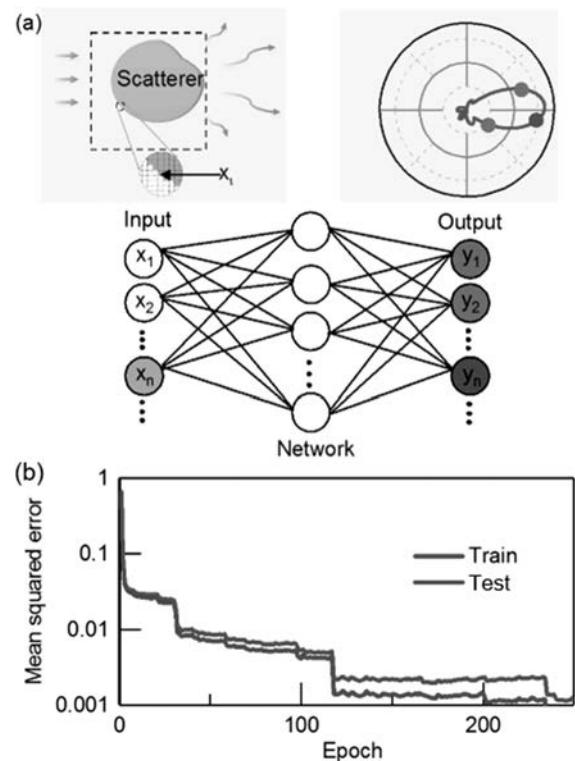
## 2. TRAINING OF DNN PREDICTORS

Here we use 2D scatterers to demonstrate our method. It can be easily generalized to high-dimensional spaces. We consider a silicon dioxide ( $\epsilon_r = 2.1$ ) scatterer in air ( $\epsilon_r = 1$ ) with a finite dimension. For example, here we limit the size of the scatterer to  $1.5\lambda$  on each side, where  $\lambda$  is the incident wavelength. Our goal is to build a tool that can instantaneously provide a far-field radiation profile when the SiO<sub>2</sub> scatterer is illuminated by a plane wave. The structure of the scatterer, which is a binary image, is used as the input of the neural network. The output of the network is a vector, which corresponds to the directivity of the radiation pattern in the angular range from 0 to 360 deg. The relationship of the output and the input can be seen in Fig. 1(a).

We first need to create training samples to train the network so that it can perform the function as described above. Each training sample consists of a pair of data: the structure and the far-field radiation pattern. The training set contains 87,000 training samples. In addition to the training set, we also create a test set that contains 11,000 pairs of structures and their far-field patterns. It will be used to test the performance of the trained neural network.

One crucial factor in evaluating the quality of the training data set is the diversity of the training samples. To increase the diversity, we use a large number of geometries with random features across different length scales. Specifically, we use three fundamental shapes in our design: a rectangle, a circle, and a triangle. We randomly vary the geometrical parameters including positions, the side length for rectangles, the radius of circles, and the side length and angle for the triangles. We also randomly combine the number of each shape. One of the structure examples is shown in the left panel in Fig. 1(a), with the blue part representing SiO<sub>2</sub>. The corresponding matrix size of the structure is  $30 \times 30$ , which corresponds to a spatial resolution of  $\lambda/20$ .

Next, we discuss the calculation of a far-field pattern in the training samples, which are done by full-wave simulations of Maxwell's equations. We use the 2D finite difference frequency domain (FDFD) method [24] to obtain the scattered field and total field of the input structure. In our case, the incident plane wave comes from the left side and propagates along the  $x$  axis with TE polarization. After getting the near-field radiation



**Fig. 1.** Illustration of our approach and the loss curve of the neural network. (a) Sketch of a scatterer and its far-field pattern. (b) Loss curve of our neural network.

pattern, we use near-field/far-field transformation method (i.e., the Stratton–Chu formula), to obtain the far-field patterns. The Stratton–Chu formula [25] in 2D is expressed as

$$\mathbf{E}_P = \sqrt{\lambda} \frac{jk}{4\pi} \mathbf{r}_0 \times \int [\mathbf{n} \times \mathbf{E} - \eta \mathbf{r}_0 \times (\mathbf{n} \times \mathbf{E})] e^{jkr_0} dS, \quad (1)$$

where  $\mathbf{E}$  and  $\mathbf{H}$  are fields on the surface  $S$  enclosing the scatterer,  $\mathbf{r}_0$  is the unit vector pointing from the origin to the field point  $P$ ,  $\mathbf{r}$  is the radius vector of the surface  $S$ ,  $\mathbf{n}$  is the unit normal to the surface  $S$ ,  $\eta$  is the impedance, approximately equaling  $377 \Omega$  in air,  $k$  is the wavenumber, and  $\mathbf{E}_P$  is the calculated far field in the direction from the origin toward point  $P$ . A circular boundary is used inside the scattered field when realizing the integral of transformation, and the far-field radiation pattern of the example is shown in Fig. 1(a), on the right panel. To get accurate near-to-far-field transformation, here we used a very high spatial resolution  $\lambda/100$  to perform the simulation. The simulations are performed at the Center for High Throughput Computing (CHTC) [26] at the University of Wisconsin–Madison. The far-field patterns are used as the ground-truth patterns that are consistent with the commercial software results.

The network is fully connected. The loss function is the L2 loss defined as

$$J = \frac{1}{2} \sum_i (r_i - o_i)^2, \quad (2)$$

where  $r_i$  is the magnitude of the electric field, and  $o_i$  is the network output value corresponding to the electric field at

different angles. Since the original values of the electric far field are about  $10^{-4}$ , the values multiplied by  $10^3$  are used to facilitate the training process. Figure 1(b) shows the loss curve with a rapid decline, reaching a very small value of test loss, which means the neural network works very well for our problems. We also noted that there are a few steep drops in the curve, and we assume that, as the training process proceeds, the loss is approaching some local minimums or saddle points in the parameter space, so the loss value gradually converges first. For the learning rate that we used, however, it could help the loss escape from those local minimums or saddle points and decrease again quickly after a few epochs. The absolute value that the loss drops at large epochs is smaller, which means the loss value is finally close to the global minimum.

Hyperparameters are chosen by a grid search, including the learning rate, batch size, and the number of layers and units. Our network architecture has six hidden layers with 8192 units in every layer. We use the activation function leaky rectified linear unit (leaky ReLU) [27] whose leaky rate is 0.2 for every hidden layer, and ReLU for the output layer. The reason for using ReLU for the output layer is that our far-field values are positive numbers without a specific range and nonlinear functions like tanh or sigmoid work only within a small range of input values, which limits the scope of the output values. An AdamOptimizer with the learning rate  $2 \times 10^{-5}$  is employed, and the batch size is 128. The input structures are flattened to a 900-by-1 vector, and the output layer has 1000 units to depict the far-field radiation pattern at different angles. The optimization process of DNNs is briefly introduced as follows. First, the parameters (i.e., weight and bias of each neural

node) of the neural network are randomly initialized. With the input structures, the output of the whole neural network is then calculated. The L2 loss is used to evaluate the performance of these parameters, and they are optimized subsequently by error back propagation [28] to minimize the loss value. By modifying the hyperparameters and repeating the training process, we could finally achieve a very low L2 loss and, at the same time, the far-field pattern prediction is good enough.

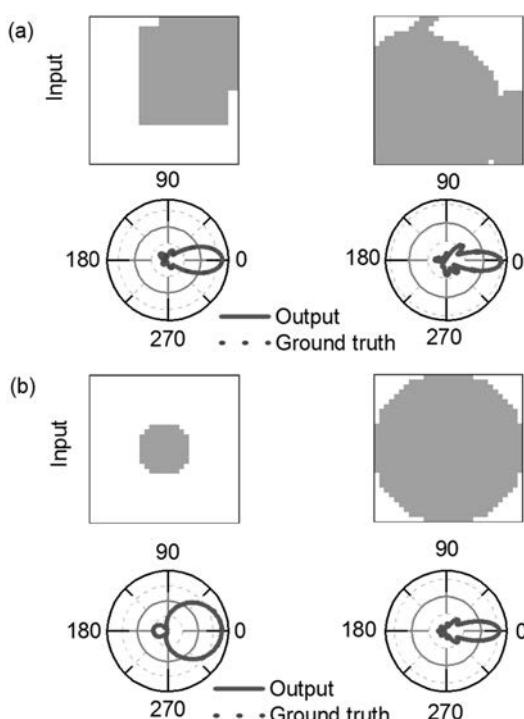
Figure 2 shows a series of results from the test set. Here we compare the far-fields patterns produced by the neural network and the ground truth, which is produced using full-wave simulations. As shown in Fig. 2(a), they agree with each other very well for a variety of different geometries. In Fig. 2(b), we show some typical scatterer shapes with different physical sizes to see if the network can handle structures with different sizes. In fact, no matter how large or small the structures are, the far-field patterns generated by the neural network are consistent with the ground truth.

On a laptop with an Intel core-i7 4720HQ, the neural network takes about 500—or even fewer—milliseconds to compute the far-field patterns. This speed makes it possible to design a far-field radiation profile in real time. One can modify the structure and instantaneously obtain real-time feedback about the far fields. Here, we further develop an online tool to demonstrate this capability.

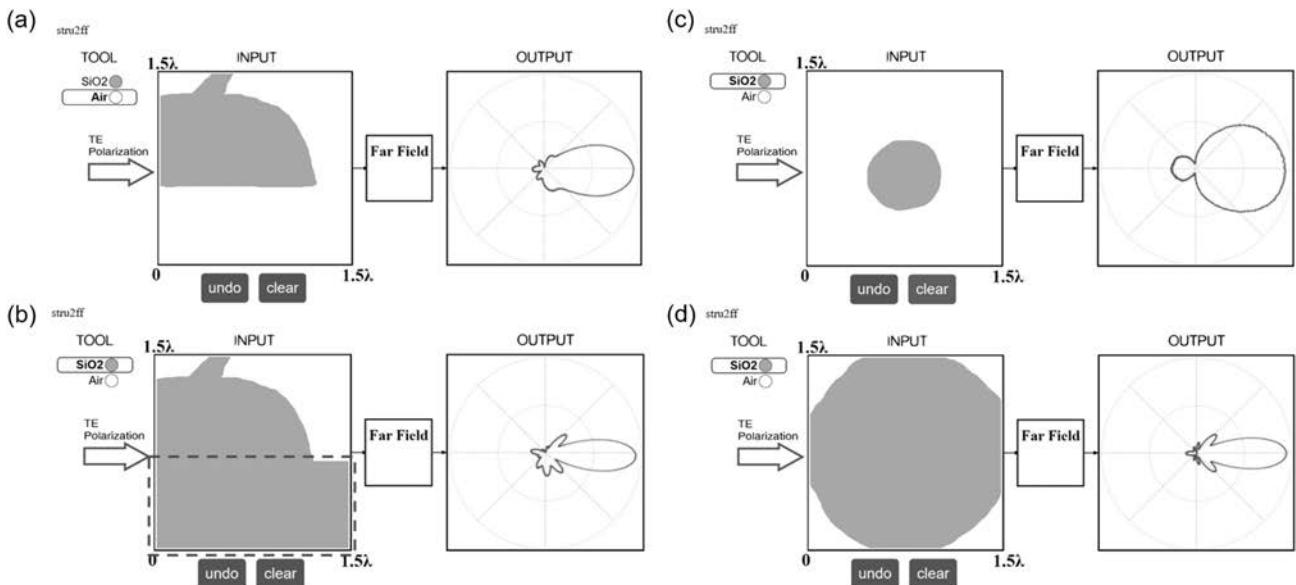
### 3. ONLINE DNN TOOL

After obtaining the trained neural network model, traditionally one needs to work with the code script to input the test structure, call the trained model, and then show the results, which is not convenient and intuitive for a practical design process. It also is not straightforward for the designers in optics without knowledge of machine learning methods to improve productivity. In this case, a method that can translate the machine learning results to an intuitive manifestation will be very useful for the optical community. Here, we design and implement a very efficient web tool to realize the calculation procedure, which combines such features as user-friendly and highly efficient to generate results, with extensibility. It also can be hosted on a website to provide easy access to a community. The main advantage of the web-page tool is that designers can intuitively use the online tool for their designs without knowing the underlying programming methods.

The application programming interface functions in TensorFlow for deep learning are transferred into the Javascript provided by Ref. [29]. Basically, we must develop two interface windows to draw the scatterer structures and exhibit the output far-field profile, and the well-trained neural network model is imported by downloading it from a known source path. For the input window, we specify two materials for the scatterer (i.e., air and silica), and one can easily extend them to multiple materials. Two functional buttons, including the undo and clear operation, are used to help the design. Then we extract the structure data from the input window as a matrix and then flatten it into a vector. The same fully connected neural network architecture as the trained model is employed with the hyperparameters and activation function to calculate the far field. The raw trained model is transferred to the binary file



**Fig. 2.** Examples of results from our method on: (a) random structures from the test set; and (b) typical shapes with different sizes, compared to the ground truth.



**Fig. 3.** Two examples of calculating a far-field pattern using our web tool. (a), (b) Two steps of drawing the final structure, which is the one on the right side in Fig. 2(a) and the related far-field pattern after each drawing stroke. (c), (d) Corresponding to the structures in Fig. 2(b). We also provide a video to show the design process online.

format for the web script to recognize the trained network parameters. Finally, we draw the far-field profile based on the polar coordinate. A screenshot of the web tool is shown in Fig. 3, and the framework is based on the one developed by Ref. [29]. In our web tool, one can hand-draw a shape on the input box, and the far-field radiation pattern will be available right away after each drawing stroke. Currently, the web page supports Firefox 67.0 or earlier. We also host the source codes on GitHub (see Code 1, Ref. [30]).

The operational details of the web are depicted as follows. There are  $256 \times 256$  pixels in the input box. During the drawing process, a handwritten figure step by step, we extract a  $30 \times 30$  matrix by downsampling from the original input matrix. At each step, we use our trained network model imported from TensorFlow to calculate and exhibit the far-field radiation pattern on the output window. Figures 3(a) and 3(b) show the operation of the tool. The far field is calculated in real time as we draw a structure. The structures we draw in Fig. 3 are similar to those in Fig. 2(a), and we can see that the far-field profile is almost the same as the ground truth. A real application process of the web tool can be found in Visualization 1. It proves that our tool will be very effective to improve the far-field design efficiency. Our tool demonstrates a new perspective to use the current emerging machine learning technologies to facilitate the complicated design process in the optics and improve productivity greatly. The neural network can also be trained to realize the inverse design [12,13,15,31,32], and thus, by using the same method, one can also integrate multiple features into the tool.

#### 4. CONCLUSION

In summary, we propose a deep learning integrated online tool to facilitate the design of far-field radiation. Unlike solving a

full-wave Maxwell's equation, our tool produces the results through DNNs in real time. Our demonstration shows that the DNN online design tool not only decreases the computation cost and time by orders of magnitude, but also provides a user-friendly platform compared to conventional software. More importantly, it shows that complicated DNNs methods could be translated to a very simple interface so that others could use it easily without any prior knowledge about a neural network. There are also other aspects of the tool that can be further improved to extend its utility, including different materials, wideband applications, and even inverse design of nanophotonics. In the future, researchers could share their trained model by integrating it with the web-based tool. We believe this will definitely enhance the usefulness of deep learning as a method to improve optical design efficiency.

**Funding.** National Natural Science Foundation of China (61974069, 62022043); Natural Science Foundation of Jiangsu Province (BK20191379); NUPTSF (NY219008); NJUPT 1311 Talent Program.

**Disclosures.** The authors declare no conflicts of interest.

#### REFERENCES

1. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature* **323**, 533–536 (1986).
2. K. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Netw.* **2**, 359–366 (1989).
3. J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities," *Proc. Natl. Acad. Sci. USA* **79**, 2554–2558 (1982).
4. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," *Appl. Opt.* **24**, 1469–1475 (1985).
5. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and

- M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
6. M. Hermans, M. Burm, T. Van Vaerenbergh, J. Dambre, and P. Bienstman, "Trainable hardware for dynamical computing using error backpropagation through physical media," *Nat. Commun.* **6**, 6729 (2015).
  7. P. Seliger, M. Mahvash, C. Wang, and A. F. J. Levi, "Optimization of aperiodic dielectric structures," *J. Appl. Phys.* **100**, 034310 (2006).
  8. A. Oskooi, A. Mutapcic, S. Noda, J. D. Joannopoulos, S. P. Boyd, and S. G. Johnson, "Robust optimization of adiabatic tapers for coupling to slow-light photonic-crystal waveguides," *Opt. Express* **20**, 21558–21575 (2012).
  9. A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vučković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nat. Photonics* **9**, 374–377 (2015).
  10. B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with  $2.4 \times 2.4 \mu\text{m}^2$  footprint," *Nat. Photonics* **9**, 378–382 (2015).
  11. P. R. Wiecha and O. L. Muskens, "Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures," *Nano Lett.* **20**, 329–338 (2019).
  12. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
  13. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
  14. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
  15. L. Gao, X. Li, D. Liu, L. Wang, and Z. Yu, "A bidirectional deep neural network for accurate silicon color design," *Adv. Mater.* **31**, 1905467 (2019).
  16. C. A. Balanis, *Antenna Theory Analysis and Design* (Wiley, 2005).
  17. P. Bharadwaj, B. Deutscher, and L. Novotny, "Optical antennas," *Adv. Opt. Photon.* **1**, 438–483 (2009).
  18. L. Novotny and N. Van Hulst, "Antennas for light," *Nat. Photonics* **5**, 83–90 (2011).
  19. N. Yu and F. Capasso, "Flat optics with designer metasurfaces," *Nat. Mater.* **13**, 139–150 (2014).
  20. A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Planar photonics with metasurfaces," *Science* **339**, 1232009 (2013).
  21. D. Lin, P. Fan, E. Hasman, and M. L. Brongersma, "Dielectric gradient metasurface optical elements," *Science* **345**, 298–302 (2014).
  22. K. Umashankar and A. Taflove, "A novel method to analyze electromagnetic scattering of complex objects," *IEEE Trans. Electromagn. Compat.* **EMC-24**, 397–405 (1982).
  23. X. Li, A. Taflove, and V. Backman, "Modified FDTD near-to-far-field transformation for improved backscattering calculation of strongly forward-scattering objects," *IEEE Antennas Wireless Propag. Lett.* **4**, 35–38 (2005).
  24. W. Shin and S. Fan, "Choice of the perfectly matched layer boundary condition for frequency-domain Maxwell's equations solvers," *J. Comput. Phys.* **231**, 3406–3431 (2012).
  25. J. A. Stratton and L. J. Chu, "Diffraction theory of electromagnetic waves," *Phys. Rev.* **56**, 99 (1939).
  26. Center for High Throughput Computing, <http://chtc.cs.wisc.edu/>.
  27. B. Xu, N. Wang, T. Chen, and M. Li, "Empirical evaluation of rectified activations in convolutional network," *arXiv:1505.00853* (2015).
  28. R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural Networks for Perception* (Academic, 1992), pp. 65–93.
  29. P. Isola, J. Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 1125–1134.
  30. <https://github.com/JRQie/Web-demo-for-far-field-pattern>.
  31. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
  32. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).

# PHOTONICS Research

## Sensing in the presence of strong noise by deep learning of dynamic multimode fiber interference

LINH V. NGUYEN,<sup>1,\*</sup> CUONG C. NGUYEN,<sup>2</sup> GUSTAVO CARNEIRO,<sup>2</sup> HEIKE EBENDORFF-HEIDEPRIEM,<sup>1,3</sup>  AND STEPHEN C. WARREN-SMITH<sup>1,3,4</sup> 

<sup>1</sup>Institute for Photonics and Advanced Sensing and School of Physical Sciences, The University of Adelaide, Adelaide, SA 5005, Australia

<sup>2</sup>Australian Institute for Machine Learning, The University of Adelaide, Adelaide, SA 5005, Australia

<sup>3</sup>Australian Research Council Centre of Excellence for Nanoscale BioPhotonics, The University of Adelaide, SA 5005, Australia

<sup>4</sup>Future Industries Institute, University of South Australia, Mawson Lakes, SA 5095, Australia

\*Corresponding author: linhnguyen.research@gmail.com

Received 25 November 2020; revised 2 February 2021; accepted 3 February 2021; posted 3 February 2021 (Doc. ID 415902); published 26 March 2021

A new approach to optical fiber sensing is proposed and demonstrated that allows for specific measurement even in the presence of strong noise from undesired environmental perturbations. A deep neural network model is trained to statistically learn the relation of the complex optical interference output from a multimode optical fiber (MMF) with respect to a measurand of interest while discriminating the noise. This technique negates the need to carefully shield against, or compensate for, undesired perturbations, as is often the case for traditional optical fiber sensors. This is achieved entirely in software without any fiber postprocessing fabrication steps or specific packaging required, such as fiber Bragg gratings or specialized coatings. The technique is highly generalizable, whereby the model can be trained to identify any measurand of interest within any noisy environment provided the measurand affects the optical path length of the MMF's guided modes. We demonstrate the approach using a sapphire crystal optical fiber for temperature sensing under strong noise induced by mechanical vibrations, showing the power of the technique not only to extract sensing information buried in strong noise but to also enable sensing using traditionally challenging exotic materials. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415902>

### 1. INTRODUCTION

Changes in the optical path length (OPL) of the guided modes in an optical fiber can be induced by various external perturbations such as temperature, strain, pressure, and biochemical binding [1–5]. This is the underlying principle of many optical fiber sensors (OFSs), particularly intrinsic type OFSs, to date. For the same reason, intrinsic OFSs are cross-sensitive to any other environmental parameters that affect the OPL beyond the intended measurand. This ubiquitous issue for intrinsic OFSs, indeed any sensor, requires careful management or compensation [3–5].

Most intrinsic OFSs developed to date are based on converting the change in an external measurand to a measurable change in OPL of the waves propagating in the fiber. Calibration of these intrinsic OFSs typically involves a single variable, where the change in the measurand is mapped to one measurable optical signal, be it a change in optical intensity or the shift of a resonant wavelength or interference fringe. Even for a spectral-based sensor such as a fiber Bragg grating (FBG), the calibration does not utilize the whole spectrum in the sense that it only focuses on the spectral shift of one

dominant optical feature, the reflected Bragg wavelength. Indeed, much of the effort in developing intrinsic OFSs is to create interference structures such as gratings or cavities so that there are such discrete resonant wavelengths. However, when changes from other perturbations are also encoded in the wavelength shift, the measurement becomes nonspecific, and such ambiguities need to be accounted for in real time. While having been implemented in various forms, the general approach to active compensation against undesired perturbations is to simultaneously measure both the intended measurand and the other competing perturbations using multiple sensor heads. Each sensor head is designed to yield different sensitivity to each measurand so that the calibration maps several optical features to each measurand to discriminate the cross-sensitivity between them [6,7]. A search with the keywords “simultaneous measurement fiber sensor” in Google Scholar returns roughly 400,000 papers for various dual-parameter sensing schemes [8].

The challenge becomes further pronounced when utilizing multimode optical fibers (MMFs) for sensing. Despite being otherwise deterministic optical systems [9], transmission through MMFs is both complex, due to the interference between many propagating optical modes, and dynamic, due

to their high sensitivity to environmental disturbances. That is, any small change in the OPL, such as through temperature, strain, bending, pressure, or biochemical binding, will lead to complex and sensitive changes in the interferometric output. Sensing can be achieved using MMFs, such as by tracking changes in the speckle output of multimode optical fibers [10], but it is often done as an ensemble measurement on the entire specklegram output rather than individual speckle features and is therefore highly susceptible to cross-sensitivity. The complex yet deterministic behavior of the MMF has also been well utilized for imaging using techniques such as optical or digital phase conjugation [11,12], wavefront optimization methods [13–17], or machine learning methods such as convolutional neural networks (CNNs) [9,18]. However, the dynamic behavior of MMFs due to undesired environmental perturbations is a hindrance to both sensing and imaging, as the propagation characteristics will change due to environmental disturbances and thus reduce image quality or cause sensor drift. It is challenging to avoid these impacts due to the complex interactions of many interfering optical modes. While deep neural networks (DNNs) have previously been applied to handle dynamic changes in the MMF, such as for spectrometry [19] and pulse characterization [20], they are fundamentally analyzing the optical input rather than quantifying the dynamic changes that are occurring to the optical fiber itself, that is, sensing.

In this work we seek to exploit both the complex and highly sensitive nature of multimode optical fibers to achieve a new paradigm in optical fiber sensing: specific measurement even in the presence of strong noise from other perturbations. Working in the wavelength domain, using a broadband input to an MMF, we vary the measurand of interest and at the same time apply random strong noise on the fiber to generate a large number of complex output spectra from the MMF. These spectra, encoded with both the measurand information and noise, are used together with the measurand labels to train a DNN to statistically relate the change of the measurand to complex changes in the output spectrum of the MMF while discriminating the noise.

The advantage of this approach is that the DNN maps the entire complex optical spectrum to the measurand change rather than focusing on single or several principal spectral features, such as an FBG peak, eliminating the need for such resonant structure for sensing. MMFs particularly present a unique platform to consider this DNN strategy due to their property of guiding many hundreds or thousands of discrete modes that can each contain important and unique information related to sensing.

We demonstrate this concept for the specific case of temperature sensing using sapphire crystal optical fiber (SOF), a single crystal fiber that has great potential for extreme temperature sensing, but where progress has been hampered by the complexity in sensing using this unstructured and thus highly multimoded waveguide [21–23]. The SOF is therefore particularly suitable for demonstrating the power of our proposed approach, but our technique can be applied, in principle, to any measurand in the presence of any environmental noise using any multimode optical waveguide.

## 2. PRINCIPLE OF OPERATION AND EXPERIMENTAL DESIGN

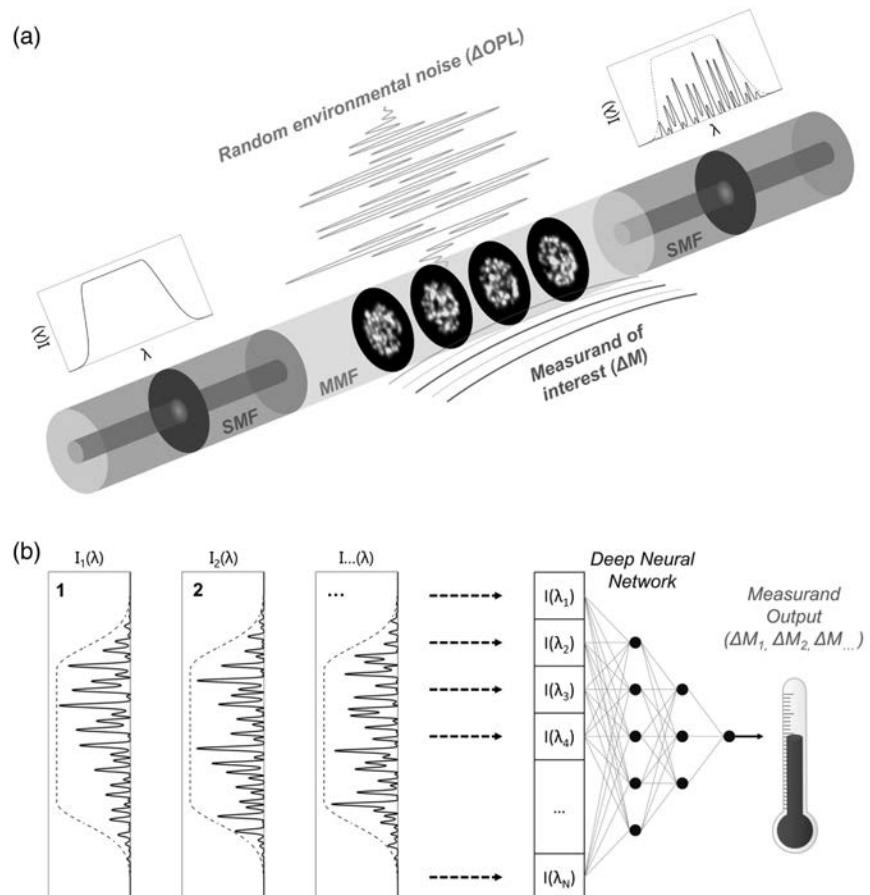
The output intensity  $I$  from an MMF, assuming coherence is maintained, is a superposition of  $N$  guided modes and can be expressed as

$$I(x, y, \lambda) = \left| \sum_{i=1}^N a_i \hat{e}_i(x, y) e^{i \frac{2\pi}{\lambda} n_i^{\text{eff}} L} \right|^2, \quad (1)$$

where  $a_i$ ,  $\hat{e}_i(x, y)$ , and  $n_i^{\text{eff}}$  are the modal amplitude, orthonormal electric field distribution, and effective refractive index of the  $i$ th mode, respectively;  $\lambda$  is the free space wavelength, and  $L$  is the length of the fiber. Equation (1) represents a speckle pattern at the output of the MMF for a given wavelength. One could measure the output speckle pattern of the MMF using a camera, but this is unsuitable for temperature sensing, the focus of our work, as imaging optics would need to be contained within a high-temperature furnace. We choose to utilize wavelength domain interference, and for this the orthogonality of the modes must be broken through spatial filtering, such as by splicing the MMF to a single-mode fiber (SMF) [24]. Through the coupling of the many modes of the MMF into the SMF, a complex interference spectrum is formed with the number of features increasing as the number of modes  $N$  increases. That is, each pair of modes in the MMF creates an interference that is transposed into a spectral interference feature after coupling into the SMF. This is equivalent to the high-dimensional data suitable for training a DNN [25].

The output interference spectrum is highly sensitive to any environmental changes on the optical fiber such as temperature, strain, and pressure, or a combination of these and others, as they affect the OPL of the multimode fiber's modes and thus the phase of the interference spectrum. For example, a temperature change will change the refractive index of the MMF material through the thermo-optic effect and thus change the effective index of the propagating modes. If the environmental changes have different impacts on different modes, albeit an intended measurand or unwanted noise, it is in principle possible to decouple these effects through analysis of the interference spectrum. In the case of temperature, each mode effective index responds differently, as the optical field has varying degrees of overlap with the MMF materials (e.g., sapphire core and air cladding), while mechanical effects such as strain largely impact the MMF length. While these changes to the output spectrum are highly sensitive and complex, in principle they contain rich information on changes that have occurred along the fiber length, and our task is to find an accurate method to extract such information.

Figure 1 shows a schematic diagram of our proposed approach for specific measurement under strong noise using MMF and DNN. The multimode interference in the wavelength domain was realized via a single-mode/multimode/single-mode fiber configuration [Fig. 1(a)]. Broadband light from a lead-in SMF is decomposed into the multiple propagation modes of the MMF. At the end of the MMF a wavelength-dependent speckle pattern is formed, which then couples to the lead-out SMF, creating a complex wavelength-dependent interference spectrum at the output. Due to the large number of modes in the MMF, this interference is highly sensitive to

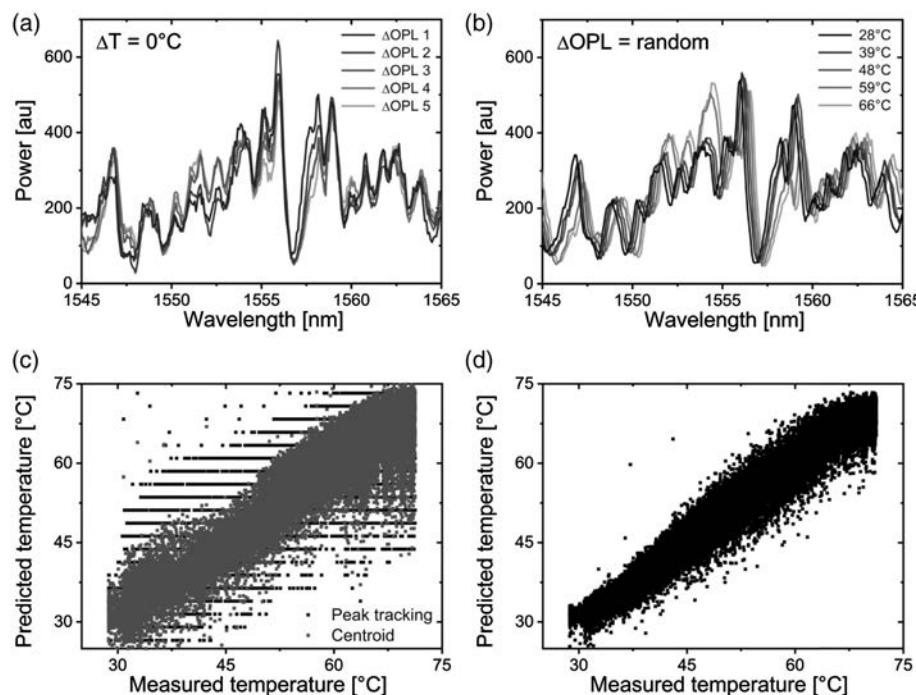


**Fig. 1.** Schematic diagram of specific measurement under strong noise using MMF and DNN. (a) An SMF-MMF-SMF structure is used to realize complex interference spectra from MMF with a broadband input. The sensing MMF is subjected to both changes in the measurand and strong noise to create complex changes in the output interference spectrum. The speckle images in the MMF conceptually show the impact of the measurand (red), which may be relatively small compared to the noise (white). (b) Various spectra together with their corresponding measurand labels are used to train a DNN. Once trained, the DNN is able to predict the measurand value buried within each interference spectrum.

the splicing condition between the fibers (e.g., tolerance in sensor fabrication) or any perturbations on the MMF. By using the DNN [Fig. 1(b)], even if the spectral change due to a chosen measurand is buried under strong spectral noise created by other environmental perturbations, it is still possible to extract the measurand information unambiguously. On the other hand, this task is extremely challenging if using traditional calibration techniques as shown in Fig. 2.

In our proof-of-concept experiment we used a 200 mm length of multimoded SOF spliced between two standard single-mode fibers (SMF28e). Details of the experimental setup are presented in Appendix A, but briefly, transmission interference spectra were measured using a swept wavelength optical sensor interrogator (National Instruments PXIe-4844) operating at 1 Hz. The spectral resolution of the interrogator was 4 pm over 1510–1590 nm, which was downsampled to 1000 spectral data points (0.8 nm spacing). The SOF was mounted loosely inside a stainless steel tube, which was vibrated with a white noise vibration spectrum using a mechanical shaker. The mounted SOF was placed into a temperature-controlled tube furnace and the temperature was recorded using a thermocouple adjacent to the SOF.

The strong noise induces significant changes in the OPL of the SOF modes as shown in Fig. 2(a). Figure 2(b) shows a similar set of spectra under the same noisy conditions at several different temperatures. The impact of the mechanical noise on the sensor makes it difficult to calibrate the wavelength shift or the intensity variation of an individual peak or trough with respect to a change in temperature. Figures 2(c) and 2(d) show the predicted temperature from the many spectra at different temperatures in the presence of strong mechanical noise using traditional calibration techniques. Figure 2(c) was obtained through direct peak tracking [26], which can be improved by applying a centroid function to determine the peak position [27]. These methods can be considered the most intuitive to apply but are not expected to perform well with complex interference spectra. A more appropriate method to extract shifts in interferometric data is to use the Fourier phase-shift method [28–31] as shown in Fig. 2(d). Here the phase of the dominant Fourier component was tracked, which corresponds to a free spectral range of 3.6 nm for the data in Fig. 1(b). It is evident that these traditional methods suffer poor accuracy in the presence of strong noise, with only 15%, 33%, and 40% of values determined within  $\pm 1^\circ\text{C}$  for the peak tracking, centroid, and



**Fig. 2.** Complex cross-sensitivity problem in specific measurement with MMF-based sensors under OPL noise. (a) Transmission interference spectra from the SOF at a fixed temperature under white noise in OPL created by shaking the loosely mounted SOF. (b) Interference spectra at several temperatures under OPL noise. (c) Predicted (calibrated) temperature by peak tracking (black) and after applying a centroid function to improve the accuracy (red). (d) Predicted (calibrated) temperature obtained using a Fourier phase-shift technique. The measured temperature in (c) and (d) was obtained using a reference thermocouple.

**Table 1. Temperature Measurement Accuracy for Different Calibration Methods<sup>a</sup>**

	$\pm 1^\circ\text{C}$	$\pm 2^\circ\text{C}$	$\pm 5^\circ\text{C}$
Peak tracking	15%	30%	70%
Centroid	33%	59%	89%
Fourier phase shift (linear)	33%	60%	96%
Fourier phase shift (quadratic)	40%	70%	97%
Deep learning	99%	100%	100%

<sup>a</sup>For the peak tracking, centroid, and Fourier methods a linear fit has been applied. In addition, a quadratic fit was also applied for the Fourier method to account for nonlinearities in the response seen in Fig. 2(d).

Fourier methods, respectively. On the other hand, as will be discussed in the following sections, a trained DNN can significantly improve the temperature measurement on the same data as shown in Table 1.

### 3. DEEP LEARNING OF THE DYNAMIC MULTIMODE INTERFERENCE

The results above confirm that without means to shield or compensate OFSs for such cross-sensitivity, calibration for the intended measurand becomes impossible. In addition, when it comes to using the multimode sensor architecture, arguably the simplest optical fiber sensing structure, any deviation in sensor preparation such as splicing and cleaving will lead to complex changes in the interference spectrum due to the involvement of many modes. These deviations are noise in their

own right. Therefore, a calibration technique that works directly with noisy data is of great advantage, as all the complexity in materials and optical engineering to mitigate cross-sensitivity is shifted to software, a much more scalable and adaptable approach.

Supervised deep learning or DNN can be thought as a highly complex fitting function that maps a high-dimensional raw input, such as an image, to an output, such as another image, the class of an object, or a measurement value [25]. The training is typically performed by showing the DNN many example inputs and requesting it to optimize a loss function, that is, the difference between ground truth and the output of the DNN model. Once the loss function is minimized for a large number of example inputs the model is considered trained and should be able to generalize to inputs it never saw during training. While DNNs have been investigated for decades [32], including early applications of DNN in MMF [33–35], recent advances in computing power, availability of data, and advanced network architectures have spawned the use of DNN for a very wide range of successful applications [25]. This includes recent promising results in reconstructing or recognizing image transmission through scattering media such as MMF [9,18,36]. In this case, the DNN is trained to understand the relation between the MMF output and input (mapping the output speckle to the original input image) where the MMF is protected from environmental perturbations. The learning capability, given a fixed network structure and data set, deteriorates with respect to the MMF length due to perturbations on the MMF such as temperature and vibration,

leading to an unstable speckle pattern at the distal end [18]. Therefore, such learning can be considered static in the sense that the transmission medium is fixed, and the training data is generated by many different input-speckle pairs.

Rather than training the DNN with many different labelled images, here we use a fixed optical input to the MMF and rely on the change of a specific measurand together with random noise along the MMF to generate a large number of different outputs, which are then used together with the measurand label to train the DNN. In this case, without prior knowledge of the light propagating in the MMF, the DNN is trained to statistically understand the relationship between the measurand and the distorted interferometric output from the MMF while ignoring the noise. In other words, the DNN is taught to learn the dynamic multimode interference with respect to an intended perturbation buried under the noise.

Figure 3 shows a schematic diagram of the DNN architecture we used for learning the relation between the transmission spectra and temperature. Here the input is MMF interference spectra inclusive of both temperature changes and noise as depicted in Fig. 1, and the DNN is trained to understand such changes buried in the OPL noise induced by strong mechanical vibration. This is analogous to the way in which DNN networks are trained to understand a cat or a dog in various lighting and background settings (noise). We used a multilayer perceptron (MLP) [37] consisting of four hidden layers, each having 512, 256, 128, and 64 nodes, respectively, resulting in 684,096 trainable parameters. The output of each layer was activated by the rectified linear unit (ReLU) [38]. No regularization, dropout, or batch normalizations were used. The loss function optimized in this regression case was the standard mean square error (MSE). An Adam optimizer [39] with a fixed learning rate of  $10^{-3}$  was used to minimize the loss. All the code was conducted using Keras—an open-source neural network written in Python [40].

While there are numerous open source deep learning architectures available, they are generally convolutional neural networks (CNNs) designed to work with 2D images that typically have a high level of spatial correlation between neighboring pixels. These large architectures are often trained using the large data set ImageNet [41], which consists of 14 million daily life

images. In contrast, the spectral data in our work is a one-dimensional array containing intensity values at different wavelengths. The wavelength spacing in our experiment was set at 0.8 nm, which we can compare to an estimate of the correlation bandwidth for a multimode optical fiber given by  $\delta\lambda \approx \lambda^2/(2\Delta nL)$ , where  $\Delta n$  is the difference between the highest and lowest effective indices supported by the optical fiber [42]. For a 20 cm length of SOF with a refractive index of 1.7 surrounded by air ( $n = 1.0$ ) the correlation bandwidth is estimated to be 8.6 pm, much less than our wavelength spacing. This means that each intensity value at one wavelength (superposition of many modes at said wavelength) is decorrelated from its neighbors and can be treated as an independent feature. Therefore, an MLP that is designed based on matrix multiplication to describe the interaction between all features within a layer is the preferred architecture for the spectral data obtained from highly multimode interference in our work. Further comparison with results obtained from training and testing the same data sets with a CNN to test this hypothesis is presented in Section 4.B.

To train and evaluate the MLP, the data set was split into three subsets: training, validation, and testing. The training and validation subsets were used to train and cross-validate to fine-tune hyperparameters such as learning rate and mini-batch size, while the testing subset was reserved for evaluation purposes.

The MLP was trained following a standard iteration process.

(i) Mini-batches of spectra were randomly sampled from the training subset and passed forward through the MLP, producing predicted temperature values.

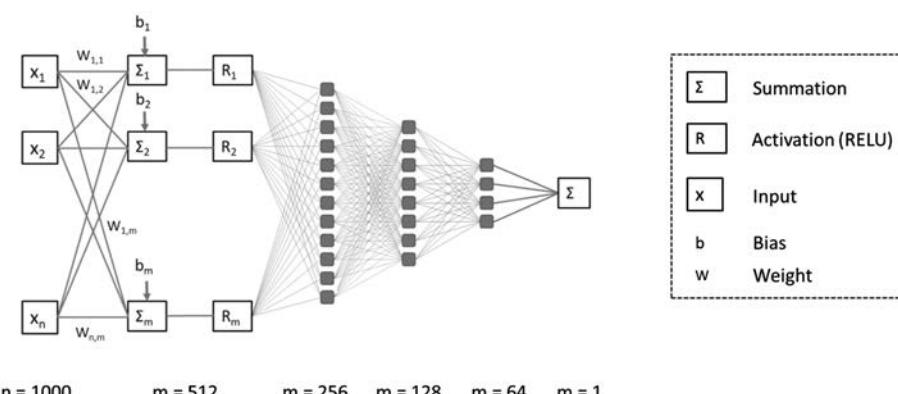
(ii) The MSE between the predicted and the actual temperatures was calculated.

(iii) The error was minimized by adjusting the network's weights following the standard backpropagation.

(iv) When all mini-batches in the training data set were completed, corresponding to 1 epoch, the training set was re-randomized and the next iteration of the training commenced.

(v) This process was repeated until the error or loss converged to a minimum value.

The training was carried out using approximately 50,000 SOF transmission spectra inclusive of random noise and their



**Fig. 3.** Schematic diagram of the multilayer perceptron (MLP) architecture used in this work. After being trained using thousands of different noisy MMF interference spectra with their corresponding temperature labels, the trained model takes an MMF spectrum as its input and produces a predicted temperature value at its output.

corresponding temperature labels. Training took approximately 7 h on an Intel Core i7-7700HQ CPU at 2.8 GHz and 32 Gb RAM until the loss converged. During the training, the prediction accuracies on the training subset and the held-out validation subset were monitored after every epoch to fine-tune hyperparameters, if necessary, to prevent overfitting.

To evaluate the trained MLP, the spectra in the testing subset were passed through the network to predict the corresponding temperatures. The obtained temperatures were then compared with the ground-truth temperatures to assess the accuracy of the predictions made by the trained MLP.

## 4. RESULTS AND DISCUSSION

### A. Temperature Sensing under Strong Noise

Figures 4(a) and 4(b) show the reference (thermocouple) measured temperatures and their corresponding predicted values from the trained model. Two different data sets were used for training as shown in Table 2. The modal accuracy is defined as the fraction of model-predicted values that are within  $\pm 1^\circ\text{C}$  or  $\pm 2^\circ\text{C}$  of the measured values. The trained MLP successfully maps almost all the values in the interference spectrum to their corresponding temperature while discriminating the effect from the OPL noise. The better accuracy for the T1 data set, reaching a  $\pm 1^\circ\text{C}$  accuracy of 99%, is attributed to the finer temperature variation in the training data compared with that of the T2 data set.

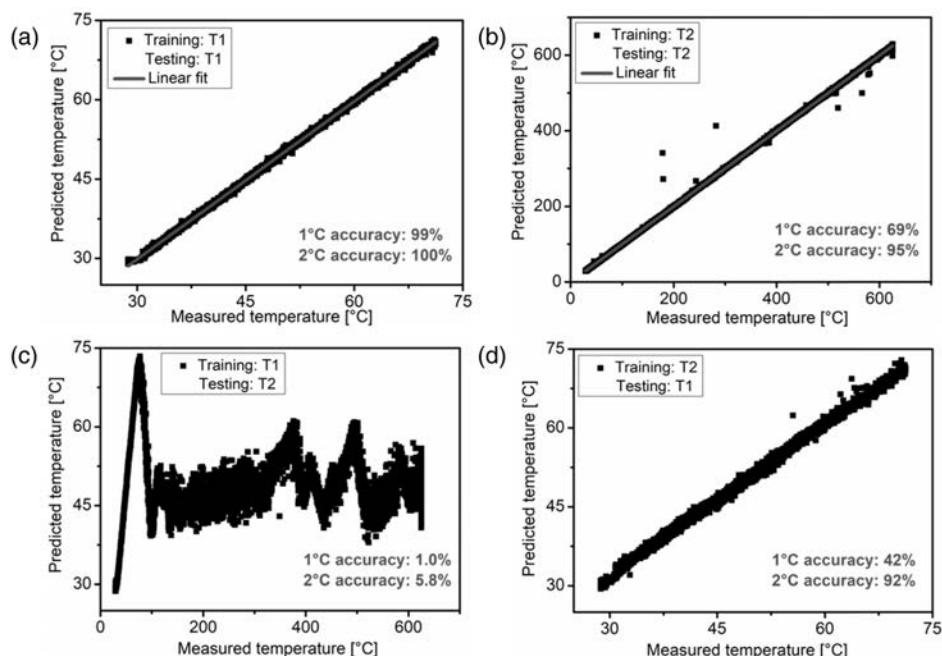
Figures 4(c) and 4(d) show the inter-testing results where the DNN is trained with one data set and tested on the other.

**Table 2. Data Set Used for Training and Testing**

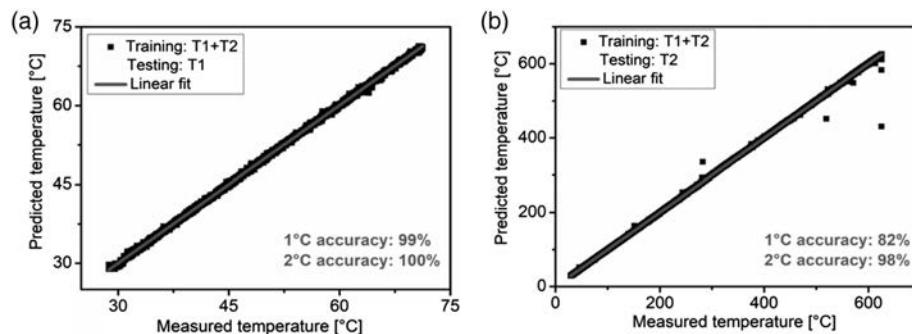
Data Set Label	Temperature Range (°C)	Ramp Rate (°C/h)	Total Spectra	Data Collection Time (h)
T1	29–75	1	84,000	23.3
T2	29–630	15	78,450	21.8

For the case of training with T1 and testing with T2, shown in Fig. 4(c), the accuracy is very poor, since there are many temperature values that are out of the training range. When the testing data T1 is within the training data T2, as shown in Fig. 4(d), the accuracy is better but still significantly less than the case of testing with the same data set T2 [e.g.,  $\pm 2^\circ\text{C}$  accuracy of 92% in Fig. 4(d) compared with 100% in Fig. 4(a)]. This is still significantly better than the Fourier shift method result (70%). We also note that the amount of training data in the temperature range of interest ( $29^\circ\text{C}$ – $75^\circ\text{C}$ ) is significantly less for Fig. 4(d) compared to Fig. 4(a) at approximately 6000 spectra compared to 78,450. This is because the  $29^\circ\text{C}$ – $75^\circ\text{C}$  region is only a small subset of the total T2 data set. One would therefore expect that the accuracy could be improved by increasing the number of spectra in the T2 training data set.

In this particular case of temperature sensing using SOF, since the furnace temperature includes OPL change along the entire SOF while the temperature labels are collected at a single point, the T1 and T2 data sets are likely shifted in their thermal distributions with respect to temperature values, and thus they would not work well in an interplay manner.



**Fig. 4.** Comparison of temperatures predicted from the trained MLP using SOF interference spectra as input and measured temperatures under OPL noise. (a) Training and testing on the same temperature range using the short temperature range from  $30^\circ\text{C}$  to  $70^\circ\text{C}$  (T1) with 8400 testing examples (data points). (b) Training and testing on the same temperature range using extended temperature range from  $30^\circ\text{C}$  to  $630^\circ\text{C}$  (T2) with 7845 testing examples. (c) Same procedure but training and testing on different data sets, the MLP is trained on the short temperature range and tested on the extended temperature range, where many temperature values are out of the training range. (d) The MLP is trained on the extended temperature range and tested on the short temperature range. In this case all the testing temperatures are predicted within the range of the training set but with less accuracy compared with (b) due to a shift in the data set.



**Fig. 5.** Mitigating the domain shift issue using a combined data set T1 + T2. (a) The testing set was T1 using 8400 example spectra. (b) The testing set was T2 using 7845 example spectra. The test results were improved for both cases.

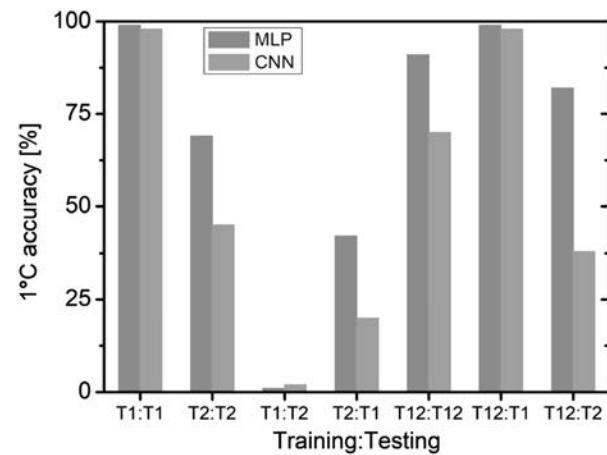
These results highlight a well-known issue of machine learning generalization where the testing data belongs to a different data distribution compared with the training data [43–46].

While mitigating the generalization issue with advanced deep learning techniques is not within the scope of this paper, readers are referred to reviews of this active area of machine learning research [47–49]. Nevertheless, the most straightforward approach is to collect more diverse data. In the case of optical fiber sensing, it implies that the noise needs to include diversity in OPL variations that mimic the changes due to many other perturbations, as well as obtaining measurand labels in a localized manner (i.e., distributed measurement) rather than a single point. Figures 5(a) and 5(b) show the improved results where the training was done on a more diverse data set, which is the combination of T1 and T2, and tested with constituent data sets. However, it is often very time consuming and expensive to have a large and diverse data set readily collected, and hence a more practical approach is to simply deploy a model trained on available data and continuously fine-tune it with newly collected data [50].

### B. Temperature Sensing: Comparison between MLP and CNN Architectures

A 1D convolutional neural network [10] with a similar number of parameters (see Appendix B) was also trained and tested using the same data sets to determine whether the model can be improved by taking into account the wavelength locality. As discussed previously in Section 3, the SOF supports many different guided modes, yielding a decorrelated interference spectrum that does not have a regular form. Therefore, it is not expected that the CNN can perform with greater accuracy compared to the MLP, in contrast to traditional image recognition problems. Figure 6 compares the 1°C accuracy between the MLP and CNN models for the cases of T1, T2, and the combined data set T1 + T2. The MLP is found to consistently perform better than its CNN counterpart for a similar number of network parameters.

The MLP and CNN architectures are two different kinds of neural networks used to model different types of data [51]. MLPs are designed based on matrix multiplication to describe the interaction between all features within a layer, allowing the network to work with any type of data. The trade-off is the linearly increasing number of parameters used. In contrast,



**Fig. 6.** Comparison between MLP and CNN architectures with a similar number of parameters. Training and testing were done in the same manner as in Figs. 4 and 5. The CNN produces inferior predictions compared with its MLP counterpart, indicating that the raw interference spectra from the highly multimode SOF do not possess significant wavelength locality.

CNNs rely on the convolution operation that models sparse interaction between features, resulting in fewer parameters. Due to the spatial nature of the convolution operation, CNNs are likely to perform better on data where the correlation is strong between adjacent features and gradually reduces for features that are further away. In our case, where the interference spectrum arises from the highly multimode SOF with many modes involved, there should be weak localized structure in the wavelength spectrum, and the superposition of all modes at a certain wavelength can be treated as an independent feature to its neighbor. Therefore, the MLP is preferable over the CNN in learning this type of data, due to both the increased accuracy shown in Fig. 6 and the relative simplicity of the network.

## 5. CONCLUSIONS AND OUTLOOK

We have proposed the use of deep learning to learn the dynamic behavior of multimode fiber interference in order to achieve specific measurement even under strong noise, without shielding the sensor from, or compensating for, cross-sensitivity

from undesired perturbations. We have achieved a proof-of-concept demonstration for the specific example of mapping the interference spectrum of an SOF and temperature, under induced OPL noise from mechanical shaking. From here our proposed technique can be expanded to any type of measurand and OPL noise provided a sufficiently diverse data set is obtained. The technique can also be readily extended to other MMF optical waveguides, again, provided a training data set is obtained for that device.

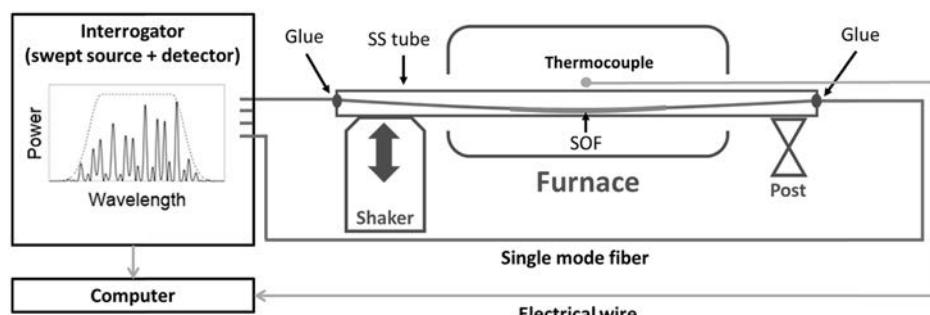
In this proof-of-concept demonstration we have focused on a single optical fiber sensor, whereby the training data and test data have been measured with the same device. A key future investigation will be to understand the generalization of our model for other nominally identical devices within prescribed tolerances. One would expect that model generalization could be achieved in a brute force approach by collecting diverse training data for many similar, but slightly different, optical fiber sensors. Understanding and solving this problem in a practical way would be key to the successful implementation of our method in an industrial setting.

This work opens an entirely new direction in optical fiber sensing in which complexity in hardware to combat cross-sensitivity or noise is shifted into software using intelligent sensor calibration that exploits advances in machine learning.

## APPENDIX A: EXPERIMENTAL SETUP

Figure 7 shows a schematic diagram of the experimental setup that was used to collect SOF interferometric data inclusive of OPL noise and temperature effect for training and testing the DNN architectures. A 200 mm length of SOF was spliced with lead-in and lead-out SMFs and inserted into a stainless steel (SS) tube of 7 mm inner diameter with both ends glued such that the SOF was loosely hung inside the SS tube. The SOF was purchased from Micromaterials and has a quoted diameter of 70  $\mu\text{m}$ .

Given the diameter and refractive index of the SOF, the fiber is estimated to support approximately 100,000 modes. However, it should be noted that the SOF is made via a crystal growth process that produces a diameter variation along the fiber, and thus the number of supported mode numbers is expected to vary along the fiber as well. The SS tube was connected to a mechanical shaker (Brüel & Kjaer LDS V455) that was driven with white noise vibration from 20 to 2000 Hz with a root-mean-squared acceleration of  $G_{\text{rms}} = 3.0\text{g}$ .



**Fig. 7.** Experimental setup. The SOF was mounted loosely inside the SS tube so that it vibrated within the SS tube. The shaker was driven by a white noise vibration profile. The SS tube was placed centrally within a tube furnace and the temperature referenced with a thermocouple.

**Table 3. Example Data Set**

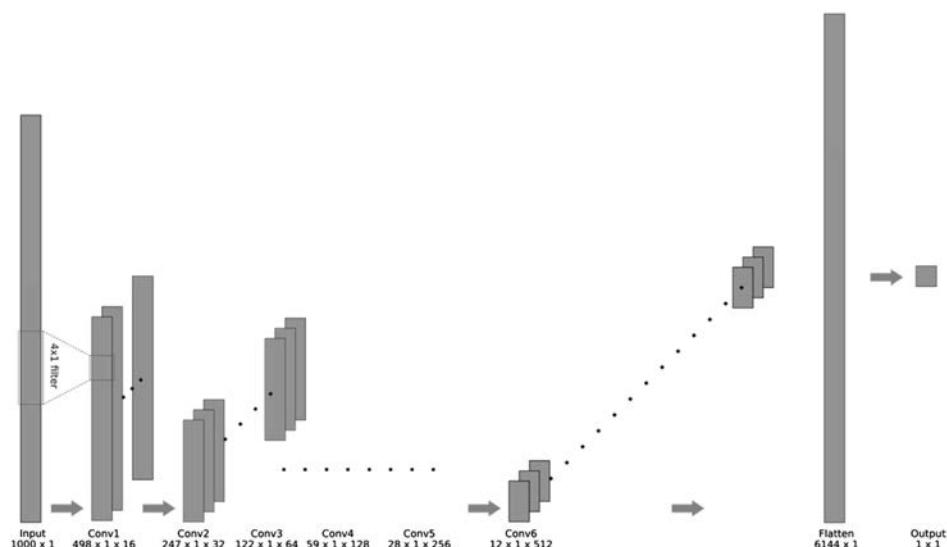
TC (°C)	$I_1$ (AU)	$I_2$ (AU)	$I_3$ (AU)	...	$I_{1000}$ (AU)
28.722	346.607	386.807	430.726	...	285.396
37.895	111.564	124.290	151.859	...	234.433
33.059	295.809	363.931	385.505	...	286.972
31.913	262.195	301.755	343.645	...	291.099
33.059	295.809	363.931	385.505	...	286.972
38.173	295.809	207.829	263.489	...	271.466
49.200	172.730	92.328	103.878	...	191.467
64.408	81.229	223.914	215.993	...	125.078
70.939	242.774	223.597	162.985	...	114.785
41.516	260.047	97.779	98.382	...	231.725

This created random movement of the SOF within the SS tube while the temperature was independently varied.

The transmission spectrum of the SMF-SOF-SMF structure was recorded using an optical sensor interrogator (National Instrument PXIe-4844). The interrogator has a wavelength resolution of 0.004 nm, which was swept at 1 Hz over a wavelength bandwidth of 80 nm so that a full range spectrum contains 20,000 intensity values. No polarization optics were used as any polarization state output from the interrogator would in any case be scrambled in the highly multimoded vibrating sapphire optical fiber. To reduce the number of dimensions in the input data, the wavelength resolution was reduced 20x to 0.08 nm, and thus the resultant spectra used for training and testing the DNN were composed of 1000 intensity values. A thermocouple was placed at the center of the SOF, and the temperature value was recorded in synchronization with the spectrum acquisition. Each data point is therefore represented as a 1001-dimensional vector whose first element is the temperature label (output) and the remaining 1000 elements are the spectrum (input). A random selection of 10 data points (spectra together with their temperature labels) is shown in Table 3. The intensity values in each spectrum were normalized to values between 0 and 1 for training.

## APPENDIX B: 1D CONVOLUTIONAL NEURAL NETWORK

A schematic diagram of the one-dimensional convolutional neural network (1D CNN) used to train the same data sets that led to the comparison with the MLP in Fig. 6 is shown in Fig. 8. It has hidden layers consisting of six convolutional



**Fig. 8.** Schematic diagram of the 1D CNN architecture.

modules and one fully connected layer with a total of 707,601 trainable parameters, comparable to that of the MLP. Each module uses a  $4 \times 1$  convolutional kernel, followed by batch normalization, ReLU, and  $2 \times 1$  max pooling. The number of channels in each layer is doubled as their sizes are halved through down-sampling. The output of the final convolutional module was flattened onto a single vector of 6144 elements and passed through the last fully connected layer to output predicted temperature. The training of the 1D CNN was similar to that of the MLP.

**Funding.** Australian Research Council (CE140100003, CE140100016, FT200100154).

**Acknowledgment.** Stephen C. Warren-Smith was funded by a Ramsay Fellowship provided by the University of Adelaide and is currently funded by an Australian Research Council Future Fellowship. This work was performed, in part, at the Optofab node of the Australian National Fabrication Facility utilizing Commonwealth and South Australian State Government Funding.

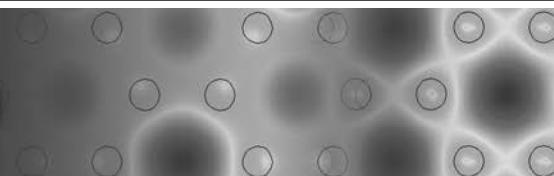
**Disclosures.** S.C.W.S. is a director of HT Sensing Pty. Ltd., a company that manufactures optical fiber sensors. HT Sensing Pty. Ltd. did not contribute to or participate in this research in any way.

## REFERENCES AND NOTES

- D. A. Krohn, T. W. MacDougall, and A. Mendez, *Fiber Optic Sensors: Fundamentals and Applications*, 4th ed. (SPIE, 2015).
- B. Lee, "Review of the present status of optical fiber sensors," *Opt. Fibre Technol.* **9**, 57–79 (2003).
- X. Wang and O. S. Wolfbeis, "Fiber-optic chemical sensors and biosensors (2015–2019)," *Anal. Chem.* **92**, 397–430 (2020).
- X. He, C. Ma, S. Wang, Z. Wang, and L. Yuan, "Pressure vector sensor based on an orthogonal optical path Sagnac interferometer," *Opt. Express* **28**, 7969–7979 (2020).
- Q. Tian, G. Xin, K.-S. Lim, Y. He, J. Liu, H. Ahmad, X. Liu, and H. Yang, "Cascaded Fabry-Perot interferometer-regenerated fiber Bragg grating structure for temperature-strain measurement under extreme temperature conditions," *Opt. Express* **28**, 30478–30488 (2020).
- V. Bhatia, D. Campbell, R. O. Claus, and A. M. Vengsarkar, "Simultaneous strain and temperature measurement with long-period gratings," *Opt. Lett.* **22**, 648–650 (1997).
- X. Li, L. V. Nguyen, H. Ebendorff-Heidepriem, D. Pham, and S. C. Warren-Smith, "Simultaneous measurement of temperature and refractive index using an exposed core micro structured optical fiber," *IEEE J. Sel. Quantum Electron.* **26**, 5600107 (2020).
- Google Scholar search using "simultaneous measurement fiber sensor," accessed October 25, 2020.
- B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," *Light Sci. Appl.* **7**, 69 (2018).
- T. D. Cabral, E. Fujiwara, S. C. Warren-Smith, H. Ebendorff-Heidepriem, and C. M. B. Cordeiro, "Multimode exposed core fiber specklegram sensor," *Opt. Lett.* **45**, 3212–3215 (2020).
- A. Yariv, "Phase conjugate optics and real-time holography," *IEEE J. Quantum Electron.* **14**, 650–660 (1978).
- I. N. Papadopoulos, S. Farahi, C. Moser, and D. Psaltis, "Focusing and scanning light through a multimode optical fiber using digital phase conjugation," *Opt. Express* **20**, 10583–10590 (2012).
- T. Čižmár and K. Dholakia, "Shaping the light transmission through a multimode optical fibre: complex transformation analysis and applications in biophotonics," *Opt. Express* **19**, 18871–18884 (2011).
- S. Bianchi and R. Di Leonardo, "A multi-mode fiber probe for holographic micromanipulation and microscopy," *Lab Chip* **12**, 635–639 (2012).
- Y. Choi, C. Yoon, M. Kim, T. D. Yang, C. Fang-Yen, R. R. Dasari, K. J. Lee, and W. Choi, "Scanner-free and wide-field endoscopic imaging by using a single multimode optical fiber," *Phys. Rev. Lett.* **109**, 203901 (2012).
- A. M. Caravaca-Aguirre, E. Niv, D. B. Conkey, and R. Piestun, "Real-time resilient focusing through a bending multimode fiber," *Opt. Express* **21**, 12881–12887 (2013).
- M. N'Gom, T. B. Norris, E. Michielssen, and R. R. Nadakuditi, "Mode control in a multimode fiber through acquiring its transmission matrix from a reference-less optical system," *Opt. Lett.* **43**, 419–422 (2018).
- N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).

19. R. K. Gupta, R. D. Bruce, S. J. Powis, and K. Dholakia, "Deep learning enabled laser speckle wavemeter with a high dynamic range," *Laser Photon. Rev.* **14**, 2000120 (2020).
20. W. Xiong, B. Redding, S. Gertler, Y. Bromberg, H. D. Tagare, and H. Cao, "Deep learning of ultrafast pulses with a multimode fiber," *APL Photon.* **5**, 096106 (2020).
21. J. Wang, B. Dong, E. Lally, J. Gong, M. Han, and A. Wang, "Multiplexed high temperature sensing with sapphire fiber air gap-based extrinsic Fabry–Perot interferometers," *Opt. Lett.* **35**, 619–621 (2010).
22. T. Habisreuther, T. Elsmann, Z. Pan, A. Graf, R. Willsch, and M. A. Schmidt, "Sapphire fiber Bragg gratings for high temperature and dynamic temperature diagnostics," *Appl. Therm. Eng.* **91**, 860–865 (2015).
23. D. Grobnić, S. J. Mihailov, C. W. Smelser, and H. Ding, "Sapphire fiber Bragg grating sensor made using femtosecond laser radiation for ultrahigh temperature applications," *IEEE Photon. Technol. Lett.* **16**, 2505–2507 (2004).
24. X. Zhu, A. Schülzgen, H. Li, L. Li, L. Han, J. V. Moloney, and N. Peyghambarian, "Detailed investigation of self-imaging in large core multimode optical fibers for application in fiber lasers and amplifiers," *Opt. Express* **16**, 16632–16645 (2008).
25. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
26. T. Wieduwilt, J. Dellith, F. Talkenberg, H. Bartelt, and M. A. Schmidt, "Reflectivity enhanced refractive index sensor based on a fiber-integrated Fabry–Perot microresonator," *Opt. Express* **22**, 25333–25346 (2014).
27. S. C. Warren-Smith, R. Kostecki, L. V. Nguyen, and T. M. Monroe, "Fabrication, splicing, Bragg grating writing, and polyelectrolyte functionalization of exposed-core microstructured optical fibers," *Opt. Express* **22**, 29493–29504 (2014).
28. L. V. Nguyen, K. Hill, S. C. Warren-Smith, and T. M. Monroe, "Interferometric-type optical biosensor based on exposed-core microstructured optical fiber," *Sens. Actuators B* **221**, 320–327 (2015).
29. S. Pevec and D. Donlagic, "High resolution, all-fiber, micro-machined sensor for simultaneous measurement of refractive index and temperature," *Opt. Express* **22**, 16241–16253 (2014).
30. J. W. Silverstone, S. McFarlane, C. P. K. Manchee, and A. Meldrum, "Ultimate resolution for refractometric sensing with whispering gallery mode microcavities," *Opt. Express* **20**, 8284–8295 (2012).
31. Y. Jiang, "Fourier transform white-light interferometry for the measurement of fiber-optic extrinsic Fabry–Pérot interferometric sensors," *IEEE Photon. Tech. Lett.* **20**, 75–77 (2008).
32. Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Backpropagation applied to handwritten zip code recognition," *Neural Comput.* **1**, 541–551 (1989).
33. S. Aisawa, K. Noguchi, and T. Matsumoto, "Remote image classification through multimode optical fiber using a neural network," *Opt. Lett.* **16**, 645–647 (1991).
34. T. Matsumoto, M. Koga, K. Noguchi, and S. Aizawa, "Proposal for neural-network applications to fiber-optic transmission," in *Proceedings of 1990 IJCNN International Joint Conference on Neural Networks* (1990), pp. 75–80.
35. R. K. Marusarz and M. R. Sayeh, "Neural network-based multimode fiber-optic information transmission," *Appl. Opt.* **40**, 219–227 (2001).
36. P. Caramazza, O. Moran, R. Murray-Smith, and D. Faccio, "Transmission of natural scene images through a multimode fibre," *Nat. Commun.* **10**, 2029 (2019).
37. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (Springer, 2009).
38. R. H. R. Hahnloser, R. Sarpeshkar, M. A. Mahowald, R. J. Douglas, and H. S. Seung, "Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit," *Nature* **405**, 947–951 (2000).
39. D. K. Diederik and J. L. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).
40. <https://keras.io/>.
41. <http://www.image-net.org/>.
42. B. Redding, S. M. Popoff, and H. Cao, "All-fiber spectrometer based on speckle pattern reconstruction," *Opt. Express* **21**, 6584–6600 (2013).
43. J. Quiñero-Candela, M. Sugiyama, A. Schwaighofer, and N. Lawrence, *Dataset Shift in Machine Learning* (MIT, 2009).
44. R. Alaiz-Rodríguez and N. Japkowicz, "Assessing the impact of changing environments on classifier performance," in *Advances in Artificial Intelligence* (Springer, 2008), pp. 13–24.
45. M. Wang and W. Deng, "Deep visual domain adaptation: a survey," *Neurocomputing* **312**, 135–153 (2018).
46. D. Tasche, "Fisher consistency for prior probability shift," *J. Mach. Learn. Res.* **18**, 1–32 (2017).
47. S. Geman, E. Bienenstock, and R. Doursat, "Neural networks and the bias/variance dilemma," *Neural Comput.* **4**, 1–58 (1992).
48. O. Bousquet and A. Elisseeff, "Stability and generalization," *J. Mach. Learn. Res.* **2**, 499–526 (2002).
49. C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in *International Conference on Learning Representation* (2017), pp. 1–15.
50. A. Ng, "Machine learning yearning: technical strategy for AI engineers in the era of deep learning," <https://www.deeplearning.ai/machine-learning-yearning/>.
51. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).

# PHOTONICS Research



## Delay-weight plasticity-based supervised learning in optical spiking neural networks

YANAN HAN,<sup>1</sup> SHUIYING XIANG,<sup>1,2,\*</sup> ZHENXING REN,<sup>1</sup> CHENTAO FU,<sup>1</sup> AIJUN WEN,<sup>1</sup> AND YUE HAO<sup>2</sup>

<sup>1</sup>State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China

<sup>2</sup>State Key Discipline Laboratory of Wide Band Gap Semiconductor Technology, School of Microelectronics, Xidian University, Xi'an 710071, China

\*Corresponding author: jxsys@126.com

Received 29 October 2020; revised 13 December 2020; accepted 3 February 2021; posted 5 February 2021 (Doc. ID 413742); published 26 March 2021

We propose a modified supervised learning algorithm for optical spiking neural networks, which introduces synaptic time-delay plasticity on the basis of traditional weight training. Delay learning is combined with the remote supervised method that is incorporated with photonic spike-timing-dependent plasticity. A spike sequence learning task implemented via the proposed algorithm is found to have better performance than via the traditional weight-based method. Moreover, the proposed algorithm is also applied to two benchmark data sets for classification. In a simple network structure with only a few optical neurons, the classification accuracy based on the delay-weight learning algorithm is significantly improved compared with weight-based learning. The introduction of delay adjusting improves the learning efficiency and performance of the algorithm, which is helpful for photonic neuromorphic computing and is also important specifically for understanding information processing in the biological brain. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.413742>

### 1. INTRODUCTION

As the improvement of traditional neural networks has gradually approached an upper limit, research focuses on neural networks with more biological reality. The spiking neural networks (SNNs), normally known as the third generation of artificial neural networks (ANNs), are more biologically plausible than previous ANNs [1] and have attracted more and more attention in recent decades [2–6]. Spikes transmitted in the biological neural networks enable the network to capture the rich dynamics of neurons and to integrate different information dimensions [3]. However, the information representation and processing manner has become a controversial issue and remains very challenging.

Rate coding is widely used in traditional SNNs; however, there is biological evidence that the precise timing of spikes also conveys information in nervous systems [7–9]. The precise timing of spikes enables higher information encoding capacity and lower power consumption [10–12], which are extremely important in information processing of the human brain. However, the exact learning mechanism still remains an open problem [13]. It has been shown that in the cerebellum and the cerebellar cortex, there exist signals that act like an instructor that helps the processing of information [14,15]. Several supervised learning algorithms have been proposed upon which specific problems that are tightly related to neural processing such as spike sequence learning and pattern recognition have been

solved successfully [16–20]. Remote supervised method (ReSuMe) is one of the supervised learning algorithms originally derived from the well-known Widrow-Hoff rule [17]. Based on photonic spike-timing-dependent plasticity (STDP) and anti-STDP rules, the synaptic weights can be adjusted to train the output neuron to fire spikes at the desired time.

Time-delayed transmission is an intrinsic feature in neural networks. Biological evidence shows that the transmission velocities in the nervous system can be modulated [21,22], for example, by changing the length and thickness of dendrites and axons [23]. The adjustability of both delay and weight of a synapse is referred to as synaptic plasticity. Delay plasticity has also been found to be helpful for the neuron in changing its firing behavior and synchronization, and it helps to understand the process of learning [24,25]. Delay selection and delay shift are two basic approaches incorporated in delay learning works [26,27]. To be specific, delay selection is to strengthen the weight of an optimal synapse among multiple subconnections, and delay shift is a more biologically plausible method training neurons to fire with coincident input spikes and with constant weight. Recently, researchers found that combined adjustment of delay and weight enhances the performance of an SNN [28–32]. In 2015, a delay learning remote supervised method (DL-ReSuMe) for spiking neurons was proposed to merge the delay shift approach and weight adjustment based on ReSuMe, by which the learning accuracy and learning speed are enhanced

[30]. In 2018, Taherkhani *et al.* proposed to appropriately train both weights and delays of excitatory and inhibitory neurons in a multilayer SNN to fire multiple desired spikes. Experimental evaluation on benchmark data sets shows that higher classification accuracy than single layer and a similar multilayer SNN can be achieved by the proposed method [31]. In 2020, Zhang *et al.* investigated the synaptic delay plasticity and proposed a novel learning method, where two representative supervised learning methods, ReSuMe and the perceptron based spiking neuron learning rule (PBSNLR), were studied and found to outperform the traditional synaptic weight learning methods [32].

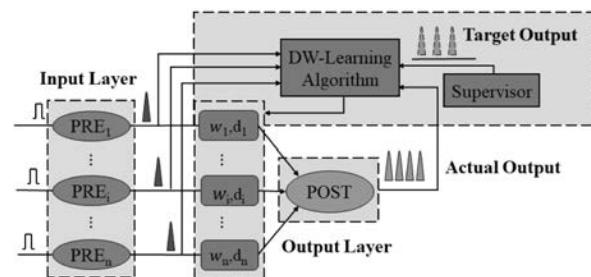
For the sake of emulating realistic biological behaviors, SNN hardware realizations are designed to seek ultralow power consumption [5]. Devices for the implementation of basic elements of SNN, namely, artificial spiking neurons and synapses, have been achieved via complementary metal-oxide-semiconductors (CMOS), transistors, and the emerging non-volatile memory technologies [33–38], etc. Photonic neuromorphic systems have attracted attention for being a potential candidate in applications of ultrafast processing. Despite its similarity with biological neurons [39], the semiconductor laser also exceeds its electronic counterpart in its ultrafast response and low power consumption. Numerous studies on photonic synapses and photonic neurons [40–53] have laid a solid foundation for significant progress in photonic neuromorphic computing based on both software and hardware implementations [54–59]. In 2019, an all-optical SNN with self-learning capacity was physically implemented on a nanophotonic chip, which is capable of supervised and unsupervised learning [55]. In 2020, we proposed to solve XOR in an all-optical neuromorphic system with inhibitory dynamics of a single photonic spiking neuron based on vertical-cavity surface-emitting lasers (VCSELs) with an embedding saturable absorber (VCSEL-SA) [58], and an all-optical spiking neural network based on VCSELs was also proposed for supervised learning and pattern classification [59]. However, as far as we know, delay learning has not yet been applied in photonic SNNs.

In this work, we propose to incorporate delay learning with the traditional algorithm based on an optical SNN. First, we propose a modified algorithm that combines delay learning with ReSuMe in a photonic SNN, which adjusts synaptic weight and delay simultaneously. By implementing a spike sequence learning task, better performance and learning efficiency of the proposed algorithm than that of its weight-based counterpart are verified. Then, the proposed algorithm is also implemented for classification, where two benchmarks, the Iris data set and the breast cancer data set, are adopted. By applying the delay-weight (DW)-based algorithm, the testing accuracy for both benchmarks is significantly improved (reaching 92%).

## 2. SYSTEM MODEL

### A. Photonic Neurons and Synapses Based on VCSEL-SA and VCSOA

The schematic diagram of DW-based supervised learning architecture is illustrated in Fig. 1, where the pre-synaptic neurons (PREs) of the input layer are fully connected to the single post-synaptic neuron (POST) in the output layer via photonic



**Fig. 1.** Schematic diagram of DW-based learning in a single-layer photonic SNN.

synapses with adjustable weight  $\omega$  and delay  $d$ . The actual output of the POST and the target output are sent to the DW learning algorithm module, based on which  $\omega$  and  $d$  can be adjusted independently to train the POST to fire spikes at a desired time. To implement a spatiotemporal encoding, the  $\text{PRE}_i$  is stimulated by a pre-coded square-wave pulse whose central timing contains time information. In a possible experiment, the modulated spikes from PREs are sent into the DW module, which contains a programmable attenuator array for the modulation of weight and a programmable true delay line (TTDL) array for the adjusting of delay. The time resolution could reach 20 ps for the TTDL devices [60], which might be sufficient for the realization of DW-ReSuMe. A possible control unit can be introduced to detect the precise spiking times of the PRE and POST photonic neurons and calculate the variation of weight and delay for each synapse during a training cycle. The training process stops when accuracy meets the requirement. Note that given the technologies available, we think it may be more realistic to adopt an *ex situ* learning approach for the training process as offline training is much easier. Once trained, the SNN can be directly used for inference implemented on optical hardware [61].

The photonic neurons and synapses are the basic elements in a photonic SNN. VCSEL-SA can mimic spiking dynamics of a biological neuron [51] and VCSEL that works below threshold value can serve as a vertical-cavity semiconductor optical amplifier (VCSOA) that is able to perform the STDP function [47], which provides possibilities of large-scale integration and low power consumption in a photonic SNN. In this work, the spiking dynamics are implemented via the excitable VCSEL-SA neurons. The rate equations of a VCSEL-SA are written as follows [57]:

$$\begin{aligned} \dot{S}_{i,o} = & \Gamma_a g_a (n_a - n_{0a}) S_{i,o} + \Gamma_s g_s (n_s - n_{0s}) S_{i,o} \\ & - S_{i,o} / \tau_{ph} + \beta B_r n_a^2, \end{aligned} \quad (1)$$

$$\begin{aligned} \dot{n}_a = & -\Gamma_a g_a (n_a - n_{0a}) (S - \Phi_{\text{pre},i} - \Phi_{\text{post},o}) \\ & - n_a / \tau_a + I_a / (eV_a), \\ \Phi_{\text{pre},i} = & k_{ei} \tau_{ph} \lambda_i P_{ei} (\tau_i, \Delta\tau) / (hcV_a), \\ \Phi_{\text{post},o} = & \sum_{i=1}^n \omega_i \lambda_i \tau_{ph} P_i (t - d_i) / (hcV_a), \end{aligned} \quad (2)$$

$$\dot{n}_s = -\Gamma_s g_s (n_s - n_{0s}) S_{i,o} - n_s / \tau_s + I_s / eV_s, \quad (3)$$

where  $i = 1, 2, \dots, n$  is the number of the PREs, and  $o$  denotes the POST. The subscripts  $a$  and  $s$  represent the gain and absorber regions, respectively.  $S_{i,o}(t)$  stands for the photon density in the cavity of the PREs and POST.  $\Gamma_a$  is the gain region confinement factor,  $g_a$  is the gain region differential gain/loss.  $n_a$  ( $n_s$ ) is the carrier density in the gain (absorber) region.

The term  $\Phi_{\text{pre},i}$  in Eq. (2) describes the pre-coded square-wave pulses injected as external stimulus into PREs, and  $\Phi_{\text{post},o}$  represents the weighted sum of all of the pre-synaptic spikes fed into the POST.  $d_i$  is the adjustable transmission delay from PRE $i$  to the PSOT.  $k_{ei}$ ,  $\tau_i$ , and  $\Delta\tau$  denote the strength, the central timing, and the temporal duration of the pulse, respectively.  $\Delta\tau_i = \tau_i - \tau_{i-1}$  is the time interval between two adjacent input pulses.  $\omega_i$  is the coupling weight between the PRE $i$  and the POST that can be tuned according to the supervised training method. The output power of PREs and POST can be calculated by  $P_{i,o}(t) \approx \eta_c \Gamma_a S_{i,o}(t) V_a b c / (\tau_{ph} \lambda_{i,o})$ . In practice, the  $\omega_i$  is calculated as an initial weight  $\omega_0$  multiplying a constant coefficient  $\omega_f = \eta_c \Gamma_a b c / (\tau_{ph} \lambda_{i,o})$  to match the optical system. The remaining parameters are the same for all neurons as in Ref. [56]. The rate equations are numerically solved by using the fourth-order Runge-Kutta method.

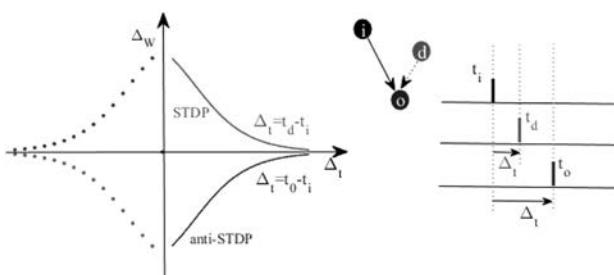
## B. DW-Based Learning Algorithm

In most of the algorithms with tasks tightly related to neural processing, such as spike sequence learning and pattern recognition, only weight adjustment is considered. However, the time delay from the PREs to the POST is hardly considered, which may play a vital role in brain computing [31]. Under this consideration, the time-delay plasticity is combined with the ReSuMe algorithm. The weight and delay changes of the  $i$ -th synapse after each training epoch are

$$\begin{aligned}\Delta\omega_i &= (n_d - n_o) + \sum_{t_d} \sum_{t_i \leq t_d} \Delta\omega_{\text{STDP}}(t_d - t_i) \\ &\quad + \sum_{t_o} \sum_{t_i \leq t_o} \Delta\omega_{\text{aSTDP}}(t_o - t_i),\end{aligned}\quad (4)$$

$$\Delta d_i = (D_{id} - D_{io}); D_{id} = t_d - t_i, D_{io} = t_o - t_i,\quad (5)$$

where  $n_d$  and  $n_o$  are spike numbers of the desired and the actual output spike sequences.  $t_i$ ,  $t_d$ , and  $t_o$  denote the input, target, and output spiking time, respectively. A schematic illustration of ReSuMe is shown in Fig. 2, where the  $\Delta\omega$  depends on three parts: namely, the non-Hebbian term, the difference between target spiking time  $t_d$  and input time  $t_i$ , and the difference



**Fig. 2.** Schematic illustration of the ReSuMe incorporated with optical STDP rule.  $i$ ,  $d$ , and  $o$  denote the input, the target, and the output, respectively.

between actual output spiking time  $t_o$  and  $t_i$ . The first term of Eq. (4) is a non-Hebbian term that aims to adjust the average strength of the input synapses to accelerate training. The  $\Delta\omega_{\text{STDP}}(t_d - t_i)$  and  $\Delta\omega_{\text{aSTDP}}(t_o - t_i)$  in Eq. (4) are photonic STDP and anti-STDP learning rules denoting the synaptic potentiation (depression), which can be calculated by [43,57]

$$\Delta\omega_{\text{STDP}}(t_d - t_i) = \begin{cases} \Delta\omega_o(\Delta t), & \text{if } t_d - t_i > 0 \\ 0, & \text{if } t_d - t_i \leq 0 \end{cases}, \quad (6)$$

$$\Delta\omega_{\text{aSTDP}}(t_o - t_i) = \begin{cases} -\Delta\omega_o(\Delta t), & \text{if } t_o - t_i > 0 \\ 0, & \text{if } t_o - t_i \leq 0 \end{cases}. \quad (7)$$

$t_i \leq t_d$  and  $t_i \leq t_o$  in Hebbian terms represent that this rule only modifies inputs that contribute to the neuron state before the desired or actual output firing time but neglect those inputs that fire afterward, which leads to better performance of the proposed DW-based algorithm.

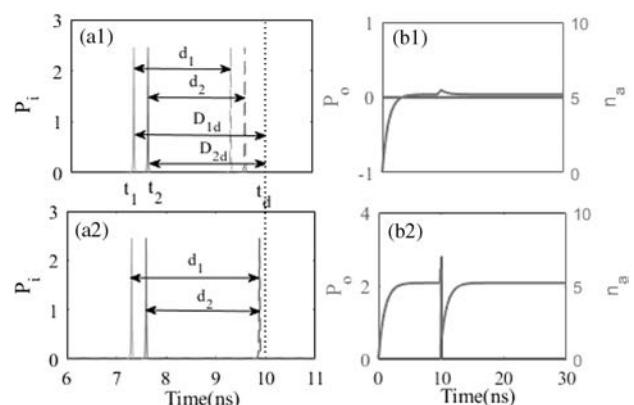
The term  $D_{id}$  ( $D_{io}$ ) is the distance between  $t_i$  and  $t_d$  ( $t_o$ ). The delay adjustment is based on the distance between input spikes and output spikes (target spikes), in a way similar to that of synaptic weight. Note that both  $\Delta\omega$  and  $\Delta d$  approach 0 if the POST fires at the desired times, which ensures the convergence of the proposed algorithm. In addition, the synaptic weights (delays) are adjusted with a learning rate  $\eta_f$  ( $\eta_d$ ) and within a learning window  $T_\omega$  ( $T_d$ ). The weight and delay of the  $i$ -th synapse are adjusted only if the input spike distance  $D_{id}$  is less than the learning window. Finally, the weight and delay of the  $i$ -th synapse are updated by

$$\omega_i(x+1) = \omega_i(x) + \eta_\omega \Delta\omega_i, \quad (8)$$

$$d_i(x+1) = d_i(x) + \eta_d \Delta d_i, \quad (9)$$

where the term  $x$  denotes the training cycle. In general,  $\eta_d = 0.5$  contributes to better performance, while other parameters should be carefully selected according to different tasks.

A simple case for delay learning is illustrated in Fig. 3. Consider two PREs; each fires a spike at  $t_1$  and  $t_2$ , respectively. After delayed transmission through the synapses, the actual input spikes locate at  $t_1 + d_1$  and  $t_2 + d_2$  [see in Fig. 3(a1)]. The principle of delay learning is to shift the actual input spikes toward the desired time, also called coherence learning, in

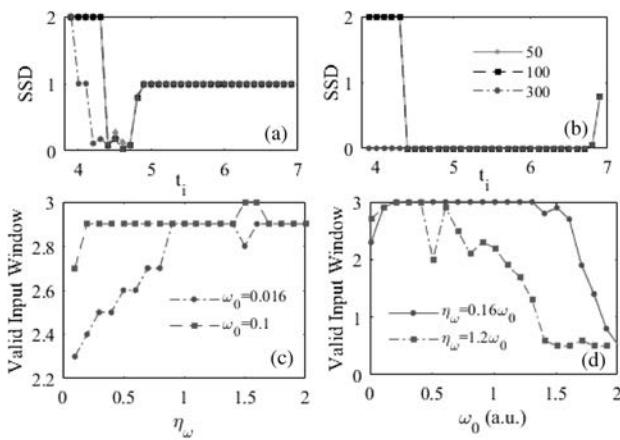


**Fig. 3.** (a1) and (b1) Input pattern and output pattern before delay adjustment. (a2) and (b2) After 7 training epochs.

which several coherent input spikes trigger the POST to fire a spike very shortly after the input time [Figs. 3(a2) and 3(b2)], while a single input spike cannot enable the release of spikes at the desired time [Fig. 3(b1)].

### 3. RESULTS AND DISCUSSION

The learning capability of a single neuron is highly related to the property of a neural network. In Fig. 4, we compare the performance of the weight-based ReSuMe and the DW-ReSuMe in recognizing a single spike, where  $N_i = N_o = 1$ ,  $d = 2$  ns, and  $t_d = 8$  ns. To quantify convergence rate, the value of spike sequence distance (SSD) (defined in Ref. [57]) is adopted. The learning window  $T_\omega$  is 3 ns, while  $T_d$  is not constrained. The input spiking time  $t_i$  varies from 4 ns to 7 ns. The time range within which an input pattern can be successfully learned, denoted as valid input window, suggests the learning ability of a neuron. The SSD at different training epochs is presented in Fig. 4 for the two algorithms. The valid input window for ReSuMe and DW-ReSuMe is 0.3 ns (4.4–4.7) and 2.4 ns (4.4–6.8) at the 50th training epoch, respectively, as shown in Figs. 4(a) and 4(b). It is shown that with DW-ReSuMe, the simple network is able to learn a single input spike with wider time range and with higher precision. Not surprisingly, with the increase of learning epochs, the valid input window also becomes a little wider. Note that weight learning is crucial when the single input is too weak to trigger an output spike in the POST. Here, we emphasize the importance of initial weight  $\omega_0$  and weight learning rate  $\eta_\omega$  on the performance of DW-based learning. When  $\omega_0$  is too small, a larger learning rate of weight is required to generate an output spike before the delay learning has shifted the input spiking time right at that of the target. As can be seen in Fig. 4(c), the efficient input window is widened by increasing the learning rate from  $0.1 \times \omega_0$  to  $0.2 \times \omega_0$ . In addition, for large  $\omega_0$ , smaller  $\eta_\omega$  is required to avoid missing the optimal solution, as indicated in Fig. 4(d). The initial weight  $\omega_0$  ranges from a unit value of



**Fig. 4.** Comparison of the learning capability of a single neuron based on (a) weight-based ReSuMe and (b) DW-ReSuMe. The value of SSD after the 50th, 100th, and 300th training epoch is presented for different  $t_i$ . (c) The valid input window as a function of  $\eta_\omega$  for different  $\omega_0$  based on DW-ReSuMe. (d) The valid input window as a function of  $\omega_0$  for different  $\eta_\omega$  based on DW-ReSuMe.  $n = 1$ ,  $t_d = 8$  ns.

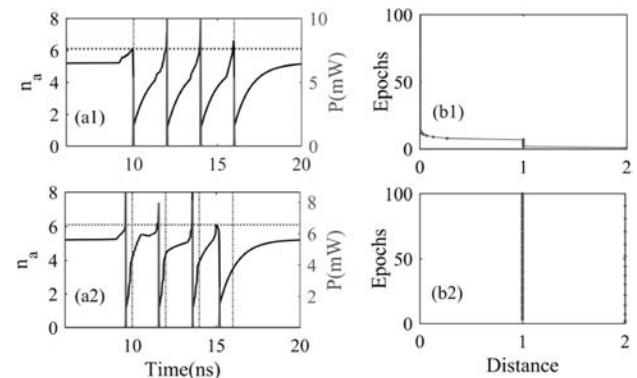
0.1 to 2, and smaller  $\eta_\omega$  contributes to a larger efficient input window. We suggest relatively larger weight learning rate  $\eta_\omega$  and smaller initial weight  $\omega_0$  for spike sequence learning. For a classification problem,  $\omega_0$  should be large enough to trigger a spike of the PREs, while smaller  $\eta_\omega$  is necessary to obtain higher accuracy.

However, note that if the desired output contains more than one spike, the delay learning window  $T_d$  has to be limited within the range of the minimum inter-spike interval (ISI) to maintain stable performance. Since some of the input spikes are shifted toward a certain target spike, there is an additional consideration in the DW-ReSuMe that the injection power of the output neuron should be limited to protect devices of photonic neurons and to ensure spiking dynamics.

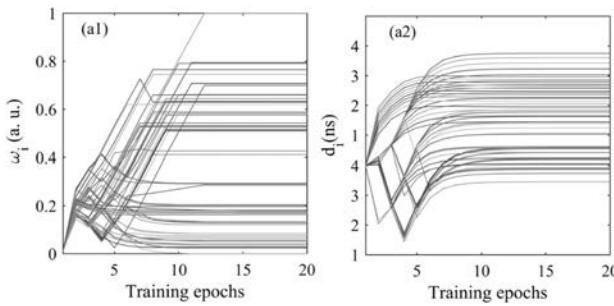
#### A. Spike Sequence Learning

Then a spike sequence learning task is implemented via a single-layer photonic SNN. Both DW-ReSuMe and ReSuMe are considered for comparison. The PRE includes 60 input neurons stimulated via pre-coded rectangular pulses with a time interval  $\Delta\tau_i$  of 0.1 ns, each connected to the POST with a photonic synapse. In Fig. 5 we show the learning process of the two algorithms. The black line in Fig. 5(a1) describes the carrier density  $n_a$  of the POST after training of 300 epochs based on DW-ReSuMe. When  $n_a$  exceeds a threshold value (marked in red dotted line), a spike is emitted. We can see that after training, the POST is able to fire accurately at the desired time (denoted by blue dotted line). The training process is further illustrated in Fig. 5(b1), where the SSD converges quickly from 2 to 0 within about 20 training epochs.

The evolution of synaptic weights and delays is shown in Fig. 6. The initial weights and delays of all neurons are identical, as indicated by Figs. 6(a) and 6(b), respectively. After training, some of the weights and delays are potentiated, while others are depressed or hardly changed. It is interesting to note that for some synapses, the weights and (or) delays change obviously at the first six training epochs. Such fluctuations during training are mainly caused by the combined effect of delay learning and weight learning. That is, when initially the output spiking time  $t_o$  is far before the target  $t_d$ , the update amount of both weight and delay is relatively large as the distance between



**Fig. 5.** (a1) Carrier density of the POST after training and (b1) the evolution of output spikes based on the DW-ReSuMe; (a2) and (b2) those based on ReSuMe. The black solid line is  $n_a$  and the red solid line represents  $P$ .



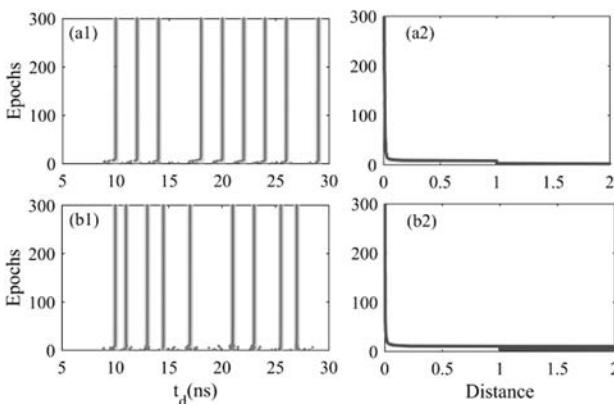
**Fig. 6.** Evolution of (a1) synaptic weights  $\omega_i$  and (a2) delays  $d_i$  during the first 20 training epochs.

$t_o$  and  $t_d$  is large. In this case, the combined adjustment of weight and delay may make  $t_o$  lag far behind  $t_d$  but closer to  $t_d$  than in the previous epoch. Note that the weight-based ReSuMe is not able to solve this problem in an SNN with 60 PREs, as shown in Figs. 5(a2) and 5(b2). The results show that combined with delay learning, DW-ReSuMe is more powerful than the weight-based algorithm.

Moreover, we found that in the cases of target spikes that have arbitrary different ISIs, the DW-ReSuMe also performs better than the traditional weight-based algorithm. Here, two cases are considered (refer to Ref. [57]). Figure 7 illustrates the learning process and evolution of SSD during 300 learning epochs. The desired spike sequence is [10 ns, 12 ns, 14 ns, 18 ns, 20 ns, 22 ns, 24 ns, 26 ns, 29 ns] in Figs. 7(a1) and 7(a2) and [10 ns, 11 ns, 13 ns, 14.5 ns, 17 ns, 21 ns, 23 ns, 25.5 ns, 27 ns] in Figs. 7(b1) and 7(b2). In both cases, with the increase of the learning epoch, the POST gradually learns to produce spikes at the desired time, and the values of SSD gradually decrease to approximate 0 after about 20 learning epochs and stay steady. Compared with our previous work [57], the network based on the DW-ReSuMe algorithm has better learning ability than that based on the weight-based ReSuMe.

### B. Fisher's Iris Data Set

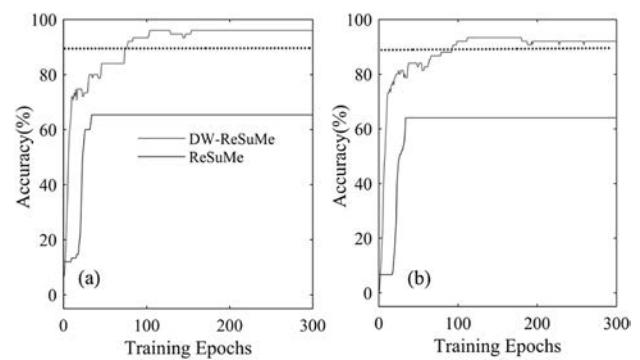
The Fisher's Iris flower data set (Fisher, 1936) is a classic benchmark of pattern recognition that contains three classes of Iris



**Fig. 7.** Learning spike sequences with ununiformed ISI. (a1) and (b1) The evolution of output spikes for spike sequence [10 ns, 12 ns, 14 ns, 18 ns, 20 ns, 22 ns, 24 ns, 26 ns, 29 ns] and [10 ns, 11 ns, 13 ns, 14.5 ns, 17 ns, 21 ns, 23 ns, 25.5 ns, 27 ns], respectively. (a2) and (b2) The evolution for the corresponding distance.

flowers with a total of 150 case entries [31]. One of the classes is linearly separable from the other two, while the other two classes are linearly inseparable. Four measurements are used to describe and differentiate the three classes, and each measurement is directly encoded into single spike firing at different times and is linearly rescaled into the interval of [5 ns, 10 ns]. The network comprises 4 PREs and a single POST. The encoding spikes are fed into the 4 PREs of the SNN via synapses with adjustable delay and weight. The initial delays are 2 ns for all of the four input neurons, and the initial weight  $\omega_0$  for each synapse is randomly selected as a constant coefficient (which is 0.1 here) multiplied by a random number from [0,1]. In this case, nearly all of the PRE neurons are capable of generating just a single spike in response of each input pattern. The weight learning rate  $\eta_\omega$  is 0.01 and decays by half every 20 training epochs to enhance the convergence of learning. The output of the network is represented by the precise spiking time of the POST, which fires a desired spike at different times for different classes. Here, the target spikes for the three classes are 8 ns, 9 ns, and 10 ns, respectively. If the output spike locates within 40% the interval of target spikes around a certain  $t_d$ , the input entry is classified into the corresponding class.

In our scheme, according to Ref. [31], 50% of the IRIS data are used for training and the rest for testing. The training and testing processes are implemented through program simulation based on MATLAB. For each training epoch, all of the entries in the training data set are injected into the input neurons of the SNN successively. Based on the learning algorithm, the  $\Delta\omega_i$  ( $\Delta d_i$ ) of the  $i$ -th input synapse for each entry is calculated and summed as  $\Delta W_i$  ( $\Delta D_i$ ). After each training epoch, the mean value of  $\Delta W_i$  ( $\Delta D_i$ ) is used for weight (delay) update, namely, the actual update amount for weight (delay) is  $\Delta\omega_i = \Delta W_i / N_{\text{train}}$  ( $\Delta d_i = \Delta D_i / N_{\text{train}}$ ), where  $N_{\text{train}} = 75$  is the size of the training data set. After each training epoch, the updated weights and delays are also used for testing. The accuracy is defined as the number of correctly classified entries divided by the total entry number. Both training accuracy and testing accuracy with different learning algorithms are presented in Fig. 8. Based on the DW-ReSuMe, the accuracy arises rapidly at first and gradually approaches 96% for the training data set

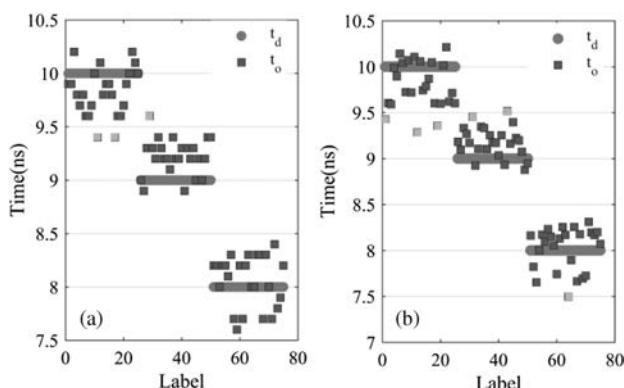


**Fig. 8.** (a) Training accuracy and (b) testing accuracy varying with training epochs for weight-based ReSuMe (blue solid line) and DW-ReSuMe (red solid line).  $T_d = 1$  ns,  $T_\omega = 4$  ns. The blue dotted line indicates an accuracy of 90%.

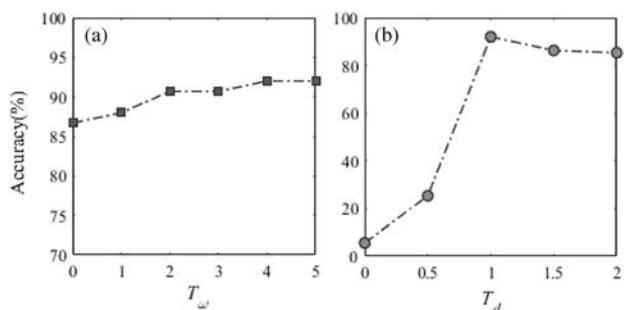
and 92% for the testing data set, as indicated by the red solid line.

For a detailed illustration of the classification results, a scattered plot of the target spiking time  $t_d$  and actual output spiking time  $t_o$  is shown in Figs. 9(a) and 9(b) for the training data set and testing data set, respectively. However, based on the same SNN architecture, the accuracy of the weight-based ReSuMe (blue solid line) only reaches 65%, with the same initial weights and delays, but a higher learning rate ( $\eta_\omega = 0.1$ ) is required to make the algorithm work efficiently. The results indicate that the performance can be greatly enhanced by introducing time-delay learning.

Moreover, we also investigate the effect of learning window on the classification accuracy. The testing accuracy as a function of weight learning window  $T_\omega$  and delay learning window  $T_d$  is shown in Figs. 10(a) and 10(b), respectively. We can see that a relatively larger  $T_\omega$  is required for achieving higher accuracy. However,  $T_d$  should be selected as the minimum ISI of target spikes, which is 1 ns. Note that the accuracy can reach more than 85% even without the adjustment of weight. The results indicate that delay learning is an extremely efficient learning algorithm in a photonic SNN with temporal encoding, which suggests that delay learning may be an essential mechanism in the biological spiking neuron systems.



**Fig. 9.** Illustration of classification results for (a) training data set and (b) testing data set. The orange cycles denote target spiking time, the blue squares represent the actual spiking time, and misclassified samples are highlighted in bright blue.



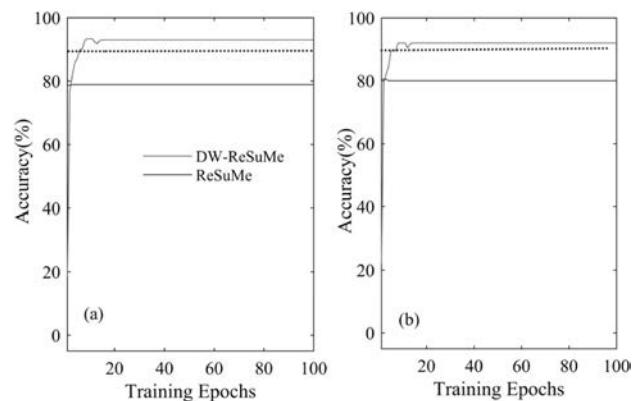
**Fig. 10.** Testing accuracy as a function of (a) weight learning window  $T_\omega$  and (b) delay learning window  $T_d$ .

### C. Breast Cancer Data Set

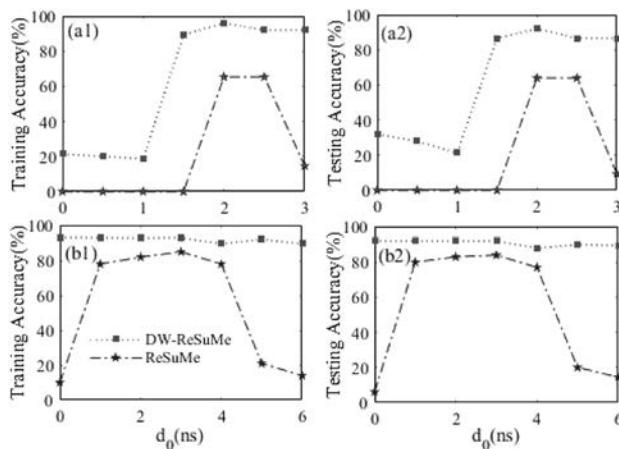
The breast cancer data set contains 608 case entries, and is divided into benign and malignant cases, each has nine measurements [62]. Five significant measurements are selected and pre-encoded into rectangular pulses, which trigger the input neurons to fire spikes at specific time within the range of 6.5 to 11 ns. The network contains 5 PREs and 1 POST. The initial delays are 0 ns for all of the 5 PREs, and the initial weights are randomly selected using the same method as for the Iris data set. The weight learning rate  $\eta_\omega$  is 0.0001. The SNN is trained to fire a spike at 9 ns for the first case and at 13 ns for the second case. Four hundred eight entries are used to train the network, with the rest for testing. The accuracy on the training data set and testing data set reaches 93% and 92%, respectively, as shown in Figs. 11(a) and 11(b). The training data set is trained for over 100 epochs, and the accuracy converges quickly within 20 training epochs. However, with weight-based ReSuMe, the classification accuracy is less than 30% with the same operating parameters (not shown here). It is worth noting that in this case, the ReSuMe algorithm does not work efficiently, limited by the simple network structure and encoding schemes. The accuracy mainly depends on the initial synaptic delay  $d_0$ . In Fig. 11, we also present the training results based on ReSuMe (blue solid line), with an initial delay of 0.5 ns. The weight learning rate here is 0.01.

Moreover, in consideration of the effect of initial delay  $d_0$  on the training performance of ReSuMe and DW-ReSuMe, we show in Fig. 12 the accuracy of the two data sets after 60 training epochs as a function of different initial delays based on the two algorithms. Both training and testing results are considered. Figures 12(a1) and 12(a2) show the training and testing accuracy of the Iris data set, respectively, and the results of the breast cancer data set are shown in Figs. 12(b1) and 12(b2). Obviously, the proposed DW-based algorithm is less sensitive to varying initial delays. Note that in the Iris data set, the learning accuracy of DW-ReSuMe is small when  $d_0$  is less than 1.5 ns, which is mainly constrained by the relatively small delay learning window (1 ns).

Finally, note that in Fig. 11 there is a downward fluctuation when the accuracy is close to the stable state. The fluctuations may be related to learning rate. A larger learning rate



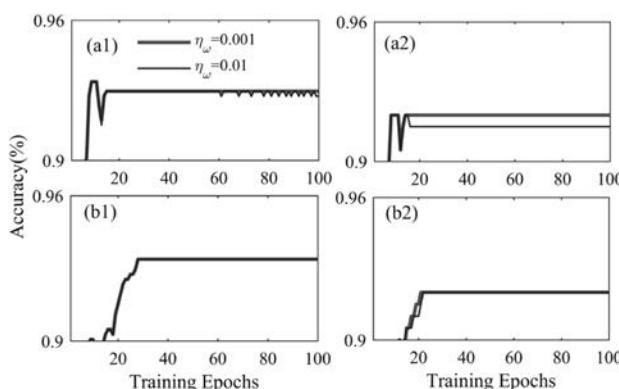
**Fig. 11.** (a) Training accuracy and (b) testing accuracy varying with training epochs based on DW-ReSuMe (red solid line) and ReSuMe (blue solid line), respectively.  $T_d = 4$  ns,  $T_\omega = 5$  ns.



**Fig. 12.** (a1) Training accuracy and (a2) testing accuracy for the Iris data set after 60 training epochs with different initial delay  $d_0$ . (b1) and (b2) The results for the breast cancer data set.

usually generates more obvious fluctuations, as illustrated in Figs. 13(a1) and 13(a2). The training accuracy fluctuates more obviously when the weight learning rate  $\eta_\omega$  is larger. However, at about the 10th training cycle, the obvious downward fluctuation is not affected by weight learning rate. We can reasonably assume that the fluctuation is caused by the relatively larger constant delay learning rate ( $\eta_d = 0.5$ ). As this value contributes to better training performance in different tasks, we consider a decaying  $\eta_d$  that is reduced by half after about five training epochs. The results are presented in Figs. 13(b1) and 13(b2), from which we can see that the fluctuation disappears. However, not surprisingly, the accuracy grows much more slowly.

DW-based learning has shown excellent performance in single-layer networks. However, the real neural networks are usually hierarchical, and synaptic weights and delays can be modulated based on different biological mechanisms, which form the foundations of complex brain functions. It is quite



**Fig. 13.** Learning accuracy of the breast cancer data set based on DW-ReSuMe with different cases of  $\eta_d$ . The left column corresponds to the training accuracy with (a1) constant  $\eta_d$  and with (b1) decaying  $\eta_d$ . (a2) and (b2) The right column shows the corresponding results of testing accuracy.  $T_d = 4$  ns,  $T_\omega = 5$  ns.

interesting and challenging to consider how to effectively introduce delay adjustment into deep learning networks.

#### 4. CONCLUSION

This paper proposed a supervised DW learning method in an optical SNN. Based on precise timing of spikes, delay learning trains coherent inputs in the input layer of an SNN via shifting the synaptic delays according to the desired and actual output timing. The proposed DW-ReSuMe is applied to spike sequence learning and two classification benchmarks, the Iris data set and breast cancer data set, successfully. The performance of the SNN is significantly improved compared with its weight-based counterpart. By introducing time-delay learning in a photonic SNN, fewer optical neurons are required to solve different tasks, which is significant to photonic neuromorphic computing. The results also suggest that synaptic delay and weight may be a combined learning mechanism in real biological neural networks, which deserves deeper investigation.

**Funding.** National Outstanding Youth Science Fund Project of National Natural Science Foundation of China (62022062); National Natural Science Foundation of China (61674119, 61974177).

**Disclosures.** The authors declare no conflicts of interest.

#### REFERENCES

- S. Ghosh-Dastidar and H. Adeli, "Spiking neural networks," *Int. J. Neural Syst.* **19**, 295–308 (2009).
- R. Brette, M. Rudolph, T. Carnevale, M. Hines, D. Beeman, J. M. Bower, M. Diesmann, A. Morrison, P. H. Goodman, F. C. Harris, Jr., M. Zirpe, T. Natschläger, D. Pecevski, B. Ermentrout, M. Djurfeldt, A. Lansner, O. Rochel, T. Vieville, E. Müller, A. P. Davison, S. El Boustani, and A. Destexhe, "Simulation of networks of spiking neurons: a review of tools and strategies," *J. Comput. Neurosci.* **23**, 349–398 (2007).
- N. Kasabov, K. Dhoble, N. Nuntalid, and G. Indiveri, "Dynamic evolving spiking neural networks for on-line spatio- and spectro-temporal pattern recognition," *Neural Netw.* **41**, 188–201 (2013).
- K. Roy, J. Akhilesh, and P. Priyadarshini, "Towards spike-based machine intelligence with neuromorphic computing," *Nature* **575**, 607–617 (2019).
- W. Q. Zhang, B. Gao, J. S. Tang, P. Yao, and H. Q. Wu, "Neuro-inspired computing chips," *Nat. Electron.* **3**, 371–382 (2020).
- T. Clarence, M. Arlja, and N. Kasabov, "Spiking neural networks: background, recent development and the NeuCube architecture," *Neural Process. Lett.* **3**, 1675–1701 (2020).
- A. Cariani, "Temporal codes and computations for sensory representation and scene analysis," *IEEE Trans. Neural Netw.* **15**, 1100–1111 (2004).
- S. M. Bohte, "The evidence for neural information processing with precise spike-times: a survey," *Natural Comput.* **3**, 195–206 (2004).
- A. Mohammed and S. Schliebs, "Training spiking neural networks to associate spatio-temporal input-output spike patterns," *Neurocomputing* **107**, 3–10 (2013).
- S. B. Laughlin, R. R. de Ruyter van Steveninck, and J. C. Anderson, "The metabolic cost of neural information," *Nat. Neurosci.* **1**, 36–41 (1998).
- S. B. Laughlin, "Energy as a constraint on the coding and processing of sensory information," *Curr. Opinion Neurobiol.* **11**, 475–480 (2001).
- H. Paugam-Moisy and S. Bohte, "Computing with spiking neuron networks," in *Handbook of Natural Computing* (Springer, 2012), pp. 335–376.

13. J. Hu, H. Tang, K. C. Tan, H. Li, and L. Shi, "A spike-timing-based integrated model for pattern recognition," *Neural Comput.* **25**, 450–472 (2013).
14. F. Ponulak and A. Kasinski, "Introduction to spiking neural networks: information processing, learning and applications," *Acta Neurobiol. Experim.* **71**, 409–433 (2011).
15. H. Jörntell and C. Hansel, "Synaptic memories upside down: bidirectional plasticity at cerebellar parallel fiber-Purkinje cell synapses," *Neuron* **52**, 227–238 (2006).
16. R. Gütig and H. Sompolinsky, "The tempotron: a neuron that learns spike timing-based decisions," *Nat. Neurosci.* **9**, 420–428 (2006).
17. F. Ponulak and A. Kasinski, "Supervised learning in spiking neural networks with ReSuMe: sequence learning, classification, and spike shifting," *Neural Comput.* **22**, 467–510 (2010).
18. I. Sporea and A. Grüning, "Supervised learning in multilayer spiking neural networks," *Neural Comput.* **25**, 473–509 (2013).
19. S. R. Kulkarni and B. Rajendran, "Spiking neural networks for handwritten digit recognition—supervised learning and network optimization," *Neural Netw.* **103**, 118–127 (2018).
20. C. Hong, X. Wei, J. Wang, B. Deng, H. Yu, and Y. Che, "Training spiking neural networks for cognitive tasks: a versatile framework compatible with various temporal codes," *IEEE Trans. Neural Netw. Learn. Syst.* **31**, 1285–1296 (2020).
21. S. Boudkkazi, L. Fronzaroli-Molinieres, and D. Debanne, "Presynaptic action potential waveform determines cortical synaptic latency," *J. Physiol.* **589**, 1117–1131 (2011).
22. J. W. Lin and D. S. Faber, "Modulation of synaptic delay during synaptic plasticity," *Trends Neurosci.* **25**, 449–455 (2002).
23. C. W. Ehrich, K. Pawelzik, U. Ernst, J. D. Cowan, and J. G. Milton, "Dynamics of self-organized delay adaptation," *Phys. Rev. Lett.* **82**, 1594–1597 (1999).
24. X. B. Gong, Y. B. Wang, and B. Ying, "Delay-induced firing behavior and transitions in adaptive neuronal networks with two types of synapses," *Sci. China Chem.* **56**, 222–229 (2012).
25. M. Dhamala, V. K. Jirsa, and M. Ding, "Enhancement of neural synchrony by time delay," *Phys. Rev. Lett.* **92**, 074104 (2004).
26. S. Ghosh-Dastidar and H. Adeli, "Improved spiking neural networks for EEG classification and epilepsy and seizure detection," *Integr. Comput.-Aided Eng.* **14**, 187–212 (2007).
27. P. Adibi, M. R. Meybodi, and R. Safabakhsh, "Unsupervised learning of synaptic delays based on learning automata in an RBF-like network of spiking neurons for data clustering," *Neurocomputing* **64**, 335–357 (2005).
28. S. Ghosh-Dastidar and H. Adeli, "A new supervised learning algorithm for multiple spiking neural networks with application in epilepsy and seizure detection," *Neural Netw.* **22**, 1419–1431 (2009).
29. A. Taherkhani, A. Belatreche, Y. Li, and L. P. Maguire, "EDL: an extended delay learning based remote supervised method for spiking neurons," in *Neural Information Processing* (Springer, 2015), pp. 190–197.
30. A. Taherkhani, A. Belatreche, Y. Li, and L. P. Maguire, "DL-ReSuMe: a delay learning-based remote supervised method for spiking neurons," *IEEE Trans. Neur. Netw. Learn. Syst.* **26**, 3137–3149 (2015).
31. A. Taherkhani, A. Belatreche, Y. Li, and L. P. Maguire, "A supervised learning algorithm for learning precise timing of multiple spikes in multilayer spiking neural networks," *IEEE Trans. Neur. Netw. Learn. Syst.* **29**, 5394–5407 (2018).
32. M. L. Zhang, J. B. Wu, A. Belatreche, Z. H. Pan, X. R. Xie, and Y. S. Chua, "Supervised learning in spiking neural networks with synaptic delay-weight plasticity," *Neurocomputing* **409**, 103–118 (2020).
33. W. Xu, S. Y. Min, H. Hwang, and T. W. Lee, "Organic core-sheath nanowire artificial synapses with femtojoule energy consumption," *Sci. Adv.* **2**, e1501326 (2016).
34. I. Sourikopoulos, S. Hedayat, C. Loyez, F. Danneville, V. Hoel, E. Mercier, and A. Cappy, "A 4-fJ/spike artificial neuron in 65 nm CMOS technology," *Front. Neurosci.* **11**, 123 (2017).
35. J. D. Zhu, Y. C. Yang, R. D. Jia, Z. X. Liang, W. Zhu, Z. A. Rehman, L. Bao, X. X. Zhang, Y. M. Cai, L. Song, and R. Huang, "Ion gated synaptic transistors based on 2D van der Waals crystals with tunable diffusive dynamics," *Adv. Mater.* **30**, 1800195 (2018).
36. B. Irem, L. G. Manuel, S. R. Nandakumar, M. Timoleon, P. Thomas, R. Bipin, L. Yusuf, S. Abu, and E. Evangelos, "Neuromorphic computing with multi-memristive synapses," *Nat. Commun.* **9**, 2514 (2018).
37. Y. Zhou, N. Xu, B. Gao, Y. Y. Chen, B. Y. Dong, Y. Li, Y. H. He, and X. S. Miao, "Complementary graphene-ferroelectric transistors (C-GFTs) as synapses with modulatable plasticity for supervised learning," in *IEEE International Electron Devices Meeting (IEDM)* (2019), pp. 1–4.
38. S. G. Hu, G. C. Qiao, Y. A. Liu, L. M. Rong, Q. Yu, and Y. Liu, "An improved memristor model connecting plastic synapse synapses and nonlinear spiking neuron," *J. Phys. D* **52**, 275402 (2019).
39. J. Ohtsubo, R. Ozawa, and M. Nanbu, "Synchrony of small nonlinear networks in chaotic semiconductor lasers," *Jpn. J. Appl. Phys.* **54**, 072702 (2015).
40. A. Hurtado, I. D. Henning, and M. J. Adams, "Optical neuron using polarization switching in a 1550 nm-VCSEL," *Opt. Express* **18**, 25170–25176 (2010).
41. A. Hurtado, K. Schires, I. Henning, and M. Adams, "Investigation of vertical cavity surface emitting laser dynamics for neuromorphic photonic systems," *Appl. Phys. Lett.* **100**, 103703 (2012).
42. M. P. Fok, Y. Tian, D. Rosenbluth, and P. R. Prucnal, "Pulse lead/lag timing detection for adaptive feedback and control based on optical spike-timing-dependent plasticity," *Opt. Lett.* **38**, 419–421 (2013).
43. Q. Li, Z. Wang, Y. Le, C. Sun, X. Song, and C. Wu, "Optical implementation of neural learning algorithms based on cross-gain modulation in a semiconductor optical amplifier," *Proc. SPIE* **10019**, 100190E (2016).
44. S. Y. Xiang, Y. H. Zhang, X. X. Guo, J. F. Li, A. J. Wen, W. Pan, and Y. Hao, "Cascadable neuron-like spiking dynamics in coupled VCSELs subject to orthogonally polarized optical pulse injection," *IEEE J. Sel. Top. Quantum Electron.* **23**, 1700207 (2017).
45. T. Deng, J. Robertson, and A. Hurtado, "Controlled propagation of spiking dynamics in vertical-cavity surface-emitting lasers: towards neuromorphic photonic networks," *IEEE J. Sel. Top. Quantum Electron.* **23**, 1800408 (2017).
46. Z. G. Chen, R. Carlos, W. H. P. Pernice, C. D. Wright, and H. Bhaskara, "On-chip photonic synapse," *Sci. Adv.* **3**, e1700160 (2017).
47. S. Y. Xiang, J. K. Gong, Y. H. Zhang, X. X. Guo, A. Wen, and Y. Hao, "Numerical implementation of wavelength-dependent photonic spike timing dependent plasticity based on VCSOA," *IEEE J. Quantum Electron.* **54**, 8100107 (2018).
48. Y. H. Zhang, S. Y. Xiang, J. K. Gong, X. X. Guo, A. J. Wen, and Y. Hao, "Spike encoding and storage properties in mutually coupled vertical-cavity surface-emitting lasers subject to optical pulse injection," *Appl. Opt.* **57**, 1731–1737 (2018).
49. I. Chakraborty, G. Saha, A. Sengupta, and K. Roy, "Toward fast neural computing using all-photonic phase change spiking neurons," *Sci. Rep.* **8**, 12980 (2018).
50. I. Chakraborty, G. Saha, and K. Roy, "Photonic in-memory computing primitive for spiking neural networks using phase-change materials," *Phys. Rev. Appl.* **11**, 014063 (2019).
51. Y. H. Zhang, S. Y. Xiang, X. Guo, A. Wen, and Y. Hao, "All-optical inhibitory dynamics in photonic neuron based on polarization mode competition in a VCSEL with an embedded saturable absorber," *Opt. Lett.* **44**, 1548–1551 (2019).
52. J. Robertson, E. Wade, Y. Kopp, J. Bueno, and A. Hurtado, "Toward neuromorphic photonic networks of ultrafast spiking laser neurons," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700715 (2020).
53. B. W. Ma and W. W. Zou, "Demonstration of a distributed feedback laser diode working as a graded-potential-signaling photonic neuron and its application to neuromorphic information processing," *Sci. China Inf. Sci.* **63**, 160408 (2020).
54. S. Y. Xiang, Y. H. Zhang, J. K. Gong, X. X. Guo, L. Lin, and Y. Hao, "STDP-based unsupervised spike pattern learning in a photonic spiking neural network with VCSELs and VCSOAs," *IEEE J. Sel. Top. Quantum Electron.* **25**, 1700109 (2019).
55. J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–215 (2019).

56. C. Ríos, N. Youngblood, Z. Cheng, M. L. Gallo, W. H. P. Pernice, C. D. Wright, A. Sebastian, and H. Bhaskaran, "In-memory computing on a photonic platform," *Sci. Adv.* **5**, eaau5759 (2019).
57. Z. W. Song, S. Y. Xiang, Z. X. Ren, G. Q. Han, and Y. Hao, "Spike sequence learning in a photonic spiking neural network consisting of VCSELs-SA with supervised training," *IEEE J. Sel. Top. Quantum Electron.* **26**, 1700209 (2020).
58. S. Y. Xiang, Z. X. Ren, Y. H. Zhang, Z. W. Song, and Y. Hao, "All-optical neuromorphic XOR operation with inhibitory dynamics of a single photonic spiking neuron based on VCSEL-SA," *Opt. Lett.* **45**, 1104–1107 (2020).
59. S. Y. Xiang, Z. X. Ren, Z. W. Song, Y. H. Zhang, X. X. Guo, G. Q. Han, and Y. Hao, "Computing primitive of fully-VCSELs-based all-optical spiking neural network for supervised learning and pattern classification," *IEEE Trans. Neural Netw.* (2020).
60. S. Moallemi, R. Welker, and J. Kitchen, "Wide band programmable true time delay block for phased array antenna applications," in *IEEE Dallas Circuits and Systems Conference (DCAS)* (2016), pp. 1–4.
61. G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, and D. Psaltis, "Inference in artificial intelligence with deep optics and photonics," *Nature* **588**, 39–47 (2020).
62. A. Roy, S. Govil, and R. Miranda, "An algorithm to generate radial basis function (RBF)-like nets for classification problems," *Neural Netw.* **8**, 179–201 (1995).

# PHOTONICS Research

## Free-space optical neural network based on thermal atomic nonlinearity

ALBERT RYOU,<sup>1,\*</sup> JAMES WHITEHEAD,<sup>1</sup> MAKSYM ZHELYEZNYAKOV,<sup>1</sup> PAUL ANDERSON,<sup>2,3</sup> CEM KESKIN,<sup>4</sup> MICHAL BAJCSY,<sup>3,5</sup> AND ARKA MAJUMDAR<sup>1,6</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, University of Washington, Seattle, Washington 98195, USA

<sup>2</sup>Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

<sup>3</sup>Institute of Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

<sup>4</sup>Google, Mountain View, California 94043, USA

<sup>5</sup>Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada

<sup>6</sup>Department of Physics, University of Washington, Seattle, Washington 98195, USA

\*Corresponding author: albertryou@gmail.com

Received 30 November 2020; revised 3 February 2021; accepted 7 February 2021; posted 8 February 2021 (Doc. ID 415964); published 26 March 2021

As artificial neural networks (ANNs) continue to make strides in wide-ranging and diverse fields of technology, the search for more efficient hardware implementations beyond conventional electronics is gaining traction. In particular, optical implementations potentially offer extraordinary gains in terms of speed and reduced energy consumption due to the intrinsic parallelism of free-space optics. At the same time, a physical nonlinearity—a crucial ingredient of an ANN—is not easy to realize in free-space optics, which restricts the potential of this platform. This problem is further exacerbated by the need to also perform the nonlinear activation in parallel for each data point to preserve the benefit of linear free-space optics. Here, we present a free-space optical ANN with diffraction-based linear weight summation and nonlinear activation enabled by the saturable absorption of thermal atoms. We demonstrate, via both simulation and experiment, image classification of handwritten digits using only a single layer and observed 6% improvement in classification accuracy due to the optical nonlinearity compared to a linear model. Our platform preserves the massive parallelism of free-space optics even with physical nonlinearity, and thus opens the way for novel designs and wider deployment of optical ANNs. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415964>

### 1. INTRODUCTION

Artificial neural networks (ANNs) have recently proven phenomenally successful in tasks such as image, sound, and language recognition and translation [1]. The increasing deployment of ANNs, from facial recognition on smartphones to self-driving cars, has brought new attention to improving their hardware implementation in terms of speed, energy consumption, and latency [2]. In contrast to conventional electronics-based platforms, optical implementations stand out due to light's intrinsically massive parallelism. For instance, the ability of a simple lens to carry out a two-dimensional (2D) Fourier transform with zero energy has long been utilized in optical signal processing [3]. Especially, free-space optics (FSO) with an aperture area  $A$  and wavelength  $\lambda$  can potentially provide an extremely large number of information channels  $\sim A/\lambda^2$ , thanks to the availability of two spatial dimensions.

One of the biggest hurdles for an optical implementation of an ANN, however, is the lack of physical optical nonlinearity.

While the parallelism of FSO naturally lends itself to carrying out linear operations, the lack of corresponding parallel nonlinearity without requiring high-powered lasers or active optical components has led to a multitude of non-FSO workaround solutions. Shen *et al.* demonstrated an electronic-optical hybrid neural network, in which the output of an integrated photonic mesh was outsourced to an external computer for nonlinear processing [4]. Nevertheless, it was shown by Colburn *et al.* that the benefits of such a design with repeated data conversions between the optical and the electronic domains were severely limited due to large power consumption and latency incurred during signal transduction [5]. Furthermore, an integrated photonics platform foregoes the intrinsic parallelism of 2D FSO. For example, Feldmann *et al.* demonstrated a fully optical spiking network with on-chip phase-change materials; but scaling up the number of neurons beyond a few waveguides remains technically challenging [6]. While wavelength division multiplexing (WDM) has been theoretically proposed as a promising route to mitigating the challenges for scaling the number of

waveguides [7], such methods need to stabilize high- $Q$  ring resonators under thermal fluctuations, leading to excess energy consumption. Moreover, a large number of additional control operations are needed to serialize the 2D image data stream and multiplex those data to encode in wavelengths, all of which will need an excess amount of energy. Another recent promising research direction is to completely avoid nonlinearity and employ multiple diffractive layers for classification [8,9]. While such an approach provided impressive classification results for the MNIST data set for a linear network combined with logistic regression, the lack of nonlinearity poses a serious question about its generalizability to solving more complicated tasks.

Recently, Zuo *et al.* presented an FSO neural network where the nonlinearity comes from the electromagnetically induced transparency in ultracold atoms [10]. Besides extensive laboratory setup for trapping and cooling atoms, the need to hand off the data from one laser to another prevents the extension of this method to having multiple hidden layers. In a similar vein, the quantum well exciton–polariton-based nonlinear activation requires a cryostat and is difficult to scale [11].

In this paper, we propose and demonstrate a fully optical ANN that utilizes the optical nonlinearity from thermal atomic vapor. Specifically, we exploit the saturable absorption behavior of room-temperature rubidium atoms housed in a vapor cell. We observed the nonlinearity in a single pass without any cavity, which allows point-by-point nonlinear activation of an incident image [12]. For the linear operations, we employ the diffractive model, where phase masks directly set the trainable weights of the neural network [8]. We emphasize that both the linear and the nonlinear components of our neural network operate on a “pixel-by-pixel” basis, within the diffraction-induced limit set by the propagation length, thus preserving and fully exploiting the intrinsic massive parallelism of FSO. Via numerical simulations, we observed an increase in classification accuracy in a single linear layer ANN by 10% due to the atomic nonlinearity. Following the training of our optical neural network in simulation using experimentally relevant parameters, we experimentally demonstrate an image recognition task of handwritten digits using a spatial light modulator (SLM). We observed an increase in classification accuracy by 6% in experiment with the addition of the nonlinear layer. We attribute the moderate classification accuracy (~33%) of our experimental system to using only a single diffractive linear operation, currently limited by the number of SLMs in our setup. Our work, combining machine learning with optics and atomic physics, opens a new front in the ongoing effort to advance optical ANN theory and hardware.

## 2. OPTICAL NEURAL NETWORK ARCHITECTURE

### A. Overview

A typical deep neural network consists of multiple layers of neurons. Except for those in the first layer, each neuron receives input signals from neurons in the previous layer. Excluding batch normalization, the neuron takes the sum of the signals multiplied by adjustable weights and performs a nonlinear operation, the output of which subsequently becomes an input signal for one or more neurons in the following layer.

Many variations in the neural network architectures exist, along with different training algorithms for specific applications. For a typical image classification task under supervised learning, the network is presented with a set of training data and corresponding labels. By repeatedly comparing the result of the output against the labels, the network can gradually adjust its weights until finally the weights converge on an optimum solution.

Our optical neural network follows a similar architecture: a 2D, monochromatic wavefront containing the input data propagates sequentially through a series of linear and nonlinear layers before being imaged on a camera. However, as explained earlier, due to a limited number of available SLMs, we only implemented one single layer that combines the input and one layer of neurons. Below, we describe each component and its physical implementation.

### B. Input Layer

The input layer is the direct representation of 2D data encoded as spatially varying intensity of light, or an image. In order to convert electronic data into optical images, we use an SLM, which can manipulate the amplitude, phase, or both of an incident laser beam’s wavefront. The use of coherent, monochromatic light is crucial for the reported optical network, since we utilize diffraction and light–matter interaction to perform both linear and nonlinear operations, as will be described next.

### C. Linear Layer

In a generic ANN, the role of a linear layer is to perform summation of signals from a previous layer with adjustable weights before passing it off to a nonlinear layer. A direct implementation of matrix–matrix multiplications in FSO is possible but complex and requires many optical elements [3]. Instead, we adopt an alternative approach, in which the linear layer is implemented by first element-wise multiplying an image with a phase mask and then letting the image propagate in free space. The first step is enabled by the SLM, which can directly display the product of an input image with the phase mask. The second step allows the signals of neighboring pixels of the image to mix due to diffraction. Such a diffractive model was demonstrated for several phase masks in the terahertz regime [8]. The amount of mixing depends on the propagation distance, the wavelength of the image, and the spatial frequency spectrum of the image. We note that it is difficult to map such a phase-mask-based approach to a traditional convolutional layer or fully connected layers used in ANNs. However, as the pixels mix with the neighboring pixels, the operation is effectively a convolution operation, the kernel of which depends on the propagation length. However, using a stack of diffractive optics [8], metasurfaces [13], or even more than one SLM, multiple layers can be implemented.

### D. Nonlinear Layer

The nonlinear layer is implemented by an evacuated vapor cell filled with rubidium atoms. The phenomenon of saturable absorption is briefly outlined here and further detailed in the Appendix. When a near-resonant photon is incident on an atom, the atom absorbs the photon and reaches an excited state. After a brief time that is inversely proportional to the atomic linewidth, the excited atom emits a photon and returns to the

ground state. The emitted photon travels in a random direction and is “lost” from the undisturbed wavefront, which continues to propagate in the original direction. Thus, for a fixed density of atoms, a low-intensity beam passing through the gas becomes attenuated. On the other hand, a high-intensity beam can excite all the available atoms, saturating the medium. The input–output curve of an optical beam of varying intensity thus exhibits a nonlinear shape, similar to the nonlinear activation function type “SmoothReLU” commonly used in machine learning. The key to our nonlinear layer is the fact that the saturation of atoms is a local effect, and thus, different parts of an incident image, which can be viewed as a collection of multiple beams with each beam denoting one pixel, undergo the nonlinear activation independently.

### E. Output Layer

The optical signal after the vapor cell is imaged on a CCD camera. The intensity pattern of the captured image becomes a direct representation of the final output of the neural network. For an image classification task with multiple categories, we can predefine certain physical locations on the camera plane to correspond to those categories. These locations then can be read by either a human or a computer to identify the categories.

We note that the absolute squaring operator inherent in taking the intensity is in itself a nonlinear process; however, as it is bound to the final measurement, we take it as part of the output layer and only refer to the independent saturable absorption layer as our nonlinearity.

## 3. SIMULATION OF A TWO-LAYER OPTICAL NEURAL NETWORK FOR IMAGE CLASSIFICATION

While atomic vapors provide a nonlinear input–output relationship, it is not clear *a priori* whether such a nonlinear function will be useful for an optical neural network, especially given that there is no energy gain in the system, only loss. To probe the efficacy of the saturable absorption nonlinearity in thermal atoms, we first simulate a two-layer optical neural network: one linear layer (to be implemented by an SLM) and one nonlinear layer (to be implemented by the saturable absorption nonlinearity). We focused on the classic image classification of handwritten digits from the MNIST database. The goal is to define an optical model, train it entirely offline, and implement the trained neural network as closely as possible in experiment. In this section, we describe the training procedure, including the use of the atomic nonlinearity, and discuss the simulation results.

The raw input data are 8-bit,  $28 \times 28$  pixels images of handwritten digits. Before feeding them to the model, we make the following modifications. First, in order that the image remains reasonably collimated during tens-of-centimeters-long free-space propagation in the experiment, we rescale the dimensions from the original  $28 \times 28$  pixels to  $300 \times 300$  pixels. Second, we further embed the  $300 \times 300$  pixels image within a larger  $600 \times 600$  pixels image, with the area outside the image having zero value. This larger dimension allows us to directly employ the angular spectrum method without applying any band limit [14]. Finally, all the values of the image-pixels are normalized so

that the maximum pixel value is 1. The size of the pixel is set to  $8 \mu\text{m}$  to match the physical pitch of our SLM. The first operation on the modified input is element-wise multiplication by a phase mask, which consists of an array of complex numbers whose magnitude is unity and whose phase  $\phi(x, y)$  is a trainable variable.

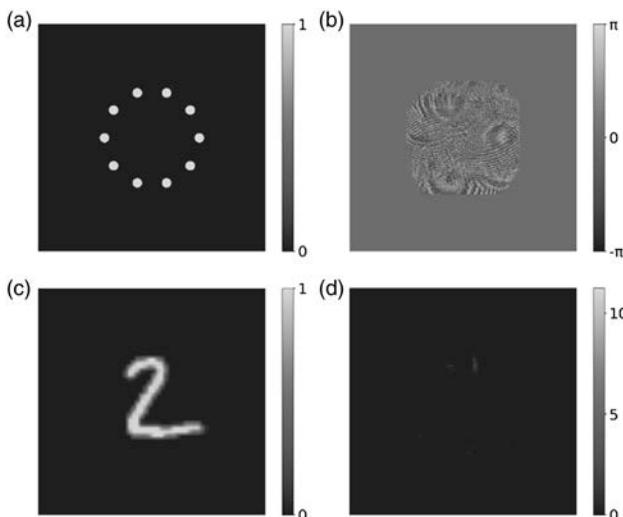
After the phase mask, the image is propagated along the optical axis by a distance  $z_o$ , which is a hyperparameter for our neural network, via the angular propagation method. The angular propagation method consists of decomposing a given wavefront into plane waves traveling in different directions, applying a  $z_o$ -dependent transfer function to each plane wave, and finally reconstructing the new wave. Computationally, the process involves a pair of forward and inverse fast Fourier transforms along with a Hadamard product with a matrix in between the pair [12].

After propagation, the image undergoes a nonlinear activation. The nonlinearity is a function of the optical intensity, so we take the absolute square of the image field, apply a nonlinear function, and take the square root, all while preserving the phase of the original wavefront. The functional form of the nonlinearity is derived in the Appendix; the nonlinear parameters were determined by a calibration process described in Section 4.B.

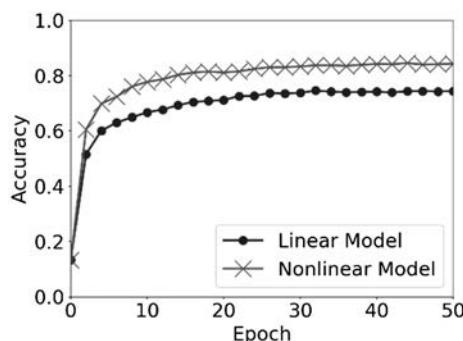
Finally, for detection, the intensity of the output of the nonlinear layer is element-wise multiplied by a special detector layer that defines where the light of a given MNIST digit should go. In our simulation, the detector layer consists of ten circles that are equidistant from the center. The result of the matrix multiplication is a list of ten numbers, each of which is the sum of the image intensity values within the circle. The maximum number, indicating the location with the highest intensity, is the final output of the neural network for the given sample image.

The entire model was defined and trained using TensorFlow 2.0 on AWS EC2. For training, we used 10,000 training images, 1000 test images, and 50 epochs, and chose Adam for optimization. The images are taken from the MNIST database. Figure 1 shows the results of a trained neural network. Figures 1(a) and 1(b) show the detector layer and the trained phase mask. The general layout of the light locations is a hyperparameter. Figures 1(c) and 1(d) show the sample input and the output of the neural network for the input. As can be seen, the location corresponding to the “2” label has the highest intensity.

Figure 2 shows the accuracy of the neural network versus the number of training epochs. In order to test the efficacy of our nonlinearity, we simulate a network without any nonlinearity (linear model) as well as the one containing the nonlinear layer (nonlinear model). As can be seen, in both cases, the accuracy rises quickly during the first few epochs and reaches a steady state. After 50 epochs, the accuracy is 74.4% for the linear model and 84.4% for the nonlinear model, showing a significant improvement. We emphasize that these accuracy values are significantly lower than that of the state-of-the-art ANNs, only because we have only one linear layer. Using multiple linear layers, our model can reach  $\sim 100\%$  accuracy for MNIST digits. However, multiple layers will be unfeasible to experimentally probe in our lab because of the limited



**Fig. 1.** Trained optical neural network (ONN). (a) The detector layer determines the location, where the light from the individual digits should be focused. The layout of the layer is a hyperparameter in our training. Here, each label corresponds to one bright circle (radius = 100  $\mu\text{m}$ ) located 1 mm from the center of the image. The “0” label is on the positive  $x$  axis, and the rest of the labels are located sequentially counterclockwise on a circle. (b) Trained phase mask; (c) sample input image; (d) output of the neural network for the sample input shown in (c). For training, the neural network calculates the intensity at each label location and returns the highest-intensity label as its prediction. All images have dimensions of 600  $\times$  600 pixels, which correspond to 4.8  $\times$  4.8 mm.



**Fig. 2.** Accuracy versus epoch for the linear model (blue dot) and the nonlinear model (red cross).

resource. Nevertheless, the increase in the classification accuracy due to the thermal atomic nonlinearity clearly shows that the physical optical nonlinearity is suitable for implementing an ANN.

## 4. EXPERIMENTAL RESULTS

### A. Setup

Figure 3 shows the layout of our experimental setup in addition to a photograph. The source of the monochromatic light is a 780-nm-wavelength laser (Toptica DL Pro), whose wavelength is fine-tuned to be resonant with the  $5S_{1/2} \rightarrow 5P_{3/2}$  transition of  $F_g = 3$   $^{85}\text{Rb}$  atoms. We first expand the collimated beam

with a 1:6 telescope in order to illuminate the SLM (Holoeye Pluto), the size of which is roughly 15.4 mm by 8.6 mm. A pair of 150-mm-focal length lenses then form a relay, between which an iris is placed to pick out the first-order diffracted beam. The resulting light, which now encodes the product of a handwritten digit and the trained phase mask, propagates for 100 mm before arriving at the front surface of a 72-mm-long vapor cell (Thorlabs GC25075-RB). The vapor cell can be lowered or raised to easily add or remove the nonlinear layer from the neural network. Furthermore, the vapor cell is wrapped in a variac-controlled heater tape (Omega) for tuning the vapor density via temperature control. During the experiment, the temperature of the cell is maintained at 50°C. Finally, a single 100-mm-focal length lens is used to image the front surface plane of the vapor cell onto a CCD camera (FLIR USB2).

### B. Nonlinearity

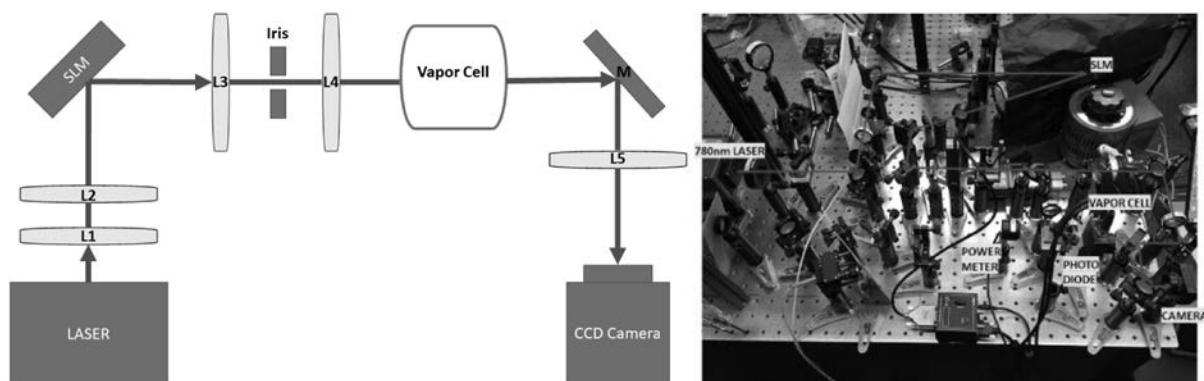
Here we describe the calibration process used to derive the nonlinear parameters for both simulation and experiment. The nonlinear input–output curve (see Appendix A) can be given by

$$I_{\text{out}} = I_{\text{in}} \exp[-N_{\text{sat}}/(1 + I_{\text{in}}/I_{\text{sat}})],$$

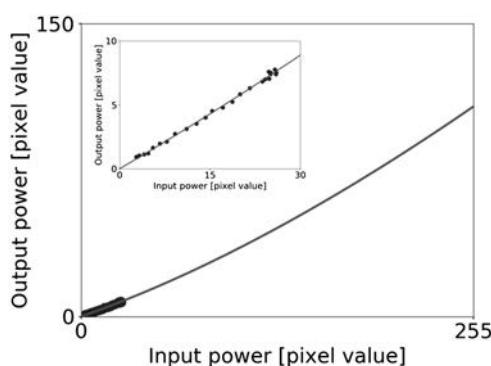
where  $I_{\text{in}}$  and  $I_{\text{out}}$  are input and output intensities,  $N_{\text{sat}}$  is the generalized atom density, and  $I_{\text{sat}}$  is the generalized saturation intensity. To determine the last two parameters, we varied the laser intensity with a wave plate and a polarizer and measured the output with and without the vapor cell. Because we are imaging the entrance plane of the vapor cell, the output measured without the vapor cell can be taken as the input into the vapor cell. From the curve fit,  $N_{\text{sat}}$  and  $I_{\text{sat}}$  were determined to be 2.6 and 0.6  $\mu\text{W}$ , respectively, which were then used for simulation in Section 3.

For experiment, it is very difficult to exactly implement the simulated model of the neural network directly due to the attenuation by many optical elements as well as the fact that the vapor cell itself has a finite length on the order of many centimeters. The latter presents a serious challenge, since the simulation assumed that the nonlinearity took place in a single plane, whereas in the experiment, the nonlinearity occurs over a continuous distance such that a propagating image would be a continuously changing attenuation.

A solution can be found if the intensity is measured not in terms of watts but the pixel value of the CCD camera itself. The parameters  $N_{\text{sat}}$  and  $I_{\text{sat}}$  then no longer refer to physical quantities, but act as general fit parameters for the nonlinear input–output curve. Figure 4 shows the plot of the average pixel value without ( $x$  axis) and with ( $y$  axis) the vapor cell in place for a sample image with varying overall intensity. The intensity range was chosen to avoid saturation of the camera without having to add an optical attenuator. Once the values of the fit parameters, 1.3 and 520, respectively, were determined, they were used to train the neural network for our experimental results with higher intensity. We emphasize that for the actual experiment, we employ the whole dynamic range of the SLM, i.e., pixel values ranging from 0 to 255.



**Fig. 3.** Experimental setup. (a) Cartoon layout of the setup. The focal lengths of the lenses are: L1, 50 mm; L2, 300 mm; L3, 150 mm; L4, 150 mm; L5, 100 mm. M indicates a flat mirror. (b) Photograph of the experiment.



**Fig. 4.** Nonlinear function showing the input–output curve for the incident intensity. The x axis is proportional to the input power, or the average pixel value on the CCD camera without the vapor cell. The y axis is proportional to the output power, or the average pixel value on the CCD camera with the vapor cell in place. Inset, zoom-in plot showing the curve fit.

### C. Results

As described before, we trained a new neural network with the nonlinear parameters that were derived directly from the camera, using 10,000 training images and 1000 test images, 100 per digit, which necessitated the adjustment of the input intensity in terms of pixel values rather than milliwatts. The resulting simulation with the experimental parameters yielded a similar-looking phase mask to that of the simulation with the ideal parameters shown in Section 3. However, the predicted accuracy dropped to 66.4% and 66.6% for the linear and the nonlinear networks, respectively, and thus, there was virtually no difference between the two networks in terms of accuracy. While it is possible in theory to achieve the original simulation regime by calibrating each optical element and reconciling simulation and experiment with more advanced techniques such as split-step nonlinear angular propagation [15], the required experimental effort and computational resources would be too great, and so we decided to proceed with the experiment.

In our experiment, we used as input the same 1000 test images that were modified as outlined in Section 3. Because our

SLM is a phase-only modulator, it cannot directly display an intensity-varying image or a complex field that is the product of the image with a phase mask; hence, we resorted to holography, which allows us to make the complex field in a conjugate plane [16] using only phase control. For detection, we calibrated the CCD camera for image magnification and rotation with separate calibration images. First, we tested the neural network that contained no phase mask. The overall accuracy was 14.7% for the linear network and 14.2% for the nonlinear network. As expected, without the phase mask, there is no significant difference between the two networks, and the accuracy is offset by a small bias near 10% (the baseline accuracy of random prediction).

Next, we repeated the test, this time incorporating the phase mask via the SLM in the neural network. The overall accuracy was measured to be 26.7% for the linear network and 33.0% for the nonlinear network. We attribute the overall reduction in accuracy compared to the simulation results to the imperfect experimental system, including fixing the length between optics, phase error in SLMs, and the finite length of the vapor cell. However, it is surprising that the accuracy is greater with incorporation of the nonlinearity, whereas the simulation shows similar performance with and without the nonlinearity. We attribute this to the robustness of the nonlinear network to the experimental noise. There is a large body of ongoing research in the machine-learning community on the effect of noise in training deep neural networks [17–19], and the exact nature of the robustness of our nonlinear optical neural network remains to be investigated. Table 1 summarizes all the accuracy results for both simulation and experiment.

**Table 1. Summary of ONN Accuracy in Percentage**

	Linear Network	Nonlinear Network
Simulation with ideal parameters	74.2	84.2
Simulation with experimental parameters	66.4	66.6
Experiment without phase mask	14.7	14.2
Experiment with phase mask	26.7	33.0

We note that the simulated and experimentally measured efficiencies are significantly lower than that of the state-of-the-art neural network. However, we have only one layer in our neural network, and we expect the accuracy to increase with a larger number of layers. Currently, our experiment is limited by available resources, i.e., a single SLM, which, while commercially available, is a significant laboratory expenditure. On the other hand, we note that creating multiple layers has several technical challenges of its own, including optical loss in each layer. The reported optical nonlinearity can be tunable by changing the temperature of the atoms, and thus can be tuned for each layer. Moreover, as we are using thermal atoms and we do not rely on cold atoms, the setup is significantly simple. However, optical regeneration techniques will be needed if the depth of the network is too large [20]. Finally, an electronic back end can be used with the optical front end to enhance the classification accuracy. We emphasize that such an electronic back end requires only one-time transduction and does not add to the overall latency, as is needed for repeated signal transduction.

#### D. Speed and Power Performance

Our reported optical ANN uses a commercial liquid crystal-based SLM with 1 million pixels, each pixel with 8-bit precision. The refresh rate of the SLM is  $\sim$ 100 Hz, making the effective supported bit rate in the optical ANN as 800 Mbps. However, using a grating light valve type of a mechanical SLM, we can increase the data rate to  $\sim$ 1 Tbps [21]. At that speed, however, we need to ensure a faster detector, e.g., an event-based camera to accommodate  $\sim\mu\text{s}$ -level detector response time [22]. Power consumption of the reported optical ANN primarily comes from the SLM, which is on the order of  $\sim$ 10 W. However, as we implement only the inference, we can use a fixed diffractive phase mask, reducing that energy to zero. For using thermal atom-based nonlinearity, we do not spend any extra energy on either active preparation or maintenance of the nonlinearity. Additionally, the reported optical ANN exploits the full potential of the parallelism offered by FSO, and thus does not require any excess energy needed for time/wavelength multiplexing. To actuate the nonlinearity, we need a light intensity of  $\sim$ 16  $\mu\text{W}/\text{mm}^2$ , and the required optical power will depend on the SLM pixel size and the optics used to guide the light through the nonlinear thermal atomic vapor. We estimate the average pixel size inside the atomic vapor to be  $\sim$ 100  $\mu\text{m} \times$  100  $\mu\text{m}$ , making the total required optical power for a million pixels  $\sim$ 160 mW. By reducing the channel area to a diffraction-limited spot ( $\sim$ 1  $\mu\text{m} \times$  1  $\mu\text{m}$ ), this power can be reduced to  $\sim$ 16  $\mu\text{W}$ .

## 5. CONCLUSION

We have shown that an atomic vapor cell can perform a local nonlinear activation in two dimensions, and consequently, a fully optical ANN can be implemented for image recognition of handwritten digits. Such a network can handle a large amount of data in parallel. Furthermore, except for the input and the output that are fed and detected by the SLM and the CCD camera, respectively, all data processing occurs in the time light takes to traverse the physical distance of the network.

Although the model accuracy of 33% is rather low, our proof-of-concept demonstrates the feasibility of using a simple, off-the-shelf atomic vapor cell as the source of fully parallel optical nonlinearity. Along with another commercially available device, the SLM, the vapor cell solves the enduring challenge of the missing optical nonlinearity that fully exploits the intrinsic massive parallelism of free-space light in two dimensions. Our work is a first step towards creating an all-optical neural network that can handle a massive amount of data and surpass the performance of an electronic neural network.

## APPENDIX A: SATURABLE ABSORPTION

Saturable absorption is a general phenomenon that appears in many different physical systems with discrete energy levels with finite lifetimes. Here we consider a simple system of two-level atoms; a detailed derivation can be found in several resources [23].

Consider a beam of photons passing through a medium with  $N$  atoms per unit volume. If the thickness of the medium is  $\Delta z$ , then the number of atoms per unit area is given by  $N\Delta z$ . If we now assign an absorption cross section  $\sigma$  to each atom, then  $N\sigma\Delta z$  is the fraction of the target area covered by the atoms. It is also the probability that an incident photon will be absorbed by the atoms, or, in the case of many photons, the total fraction of photons that are absorbed. The change in the beam intensity is then  $\Delta I/I = -N\sigma\Delta z$ , which, upon integration, yields Beer's law:  $I(z) = I_0 e^{-\kappa z}$ , where the absorption coefficient  $\kappa = N\sigma$ .

If we assume that the atoms have two levels, the ground state and the excited state, then the absorption is given by  $\kappa = (N_g - N_e)\sigma$ . Imposing the conservation of atom number ( $N_{\text{total}} = N_g + N_e$ ) and the conservation of energy [ $(N_g - N_e)\sigma I = N_e A \hbar\omega$ ] where the spontaneous decay rate  $A = 1/\tau$ , we arrive at the steady-state population difference:  $N_g - N_e = N_{\text{total}}/(1 + I/I_s)$ , where we have defined the saturation intensity  $I_{\text{sat}} = \hbar\omega A/(2\sigma)$ .

Thus, the output intensity as a function of the input intensity is given by  $I_{\text{out}} = I_{\text{in}} \exp[-N_{\text{total}}\sigma L/(1 + I_{\text{in}}/I_{\text{sat}})]$ , where  $L$  is the effective interaction length of the vapor cell. We use  $N_{\text{sat}} = N_{\text{total}}\sigma L$  and  $I_{\text{sat}}$  as our nonlinear parameters. We note that for the atomic vapor system, the variable  $N_{\text{total}}$  can be controlled by changing the temperature of the cell. Thus, the demonstrated nonlinearity is tunable, which can be exploited for a multilayer optical neural network, where the  $N_{\text{sat}}$  value will be gradually decreased to accommodate the signal loss in each layer.

**Funding.** Washington Research Foundation; UW Reality Lab; Futurewei; Facebook; Amazon; Google.

**Acknowledgment.** A. R. acknowledges support from the IC Fellowship. M. Z. acknowledges support from the NSF Graduate Research Fellowship. P. A. acknowledges support from Canada First Excellence Research Fund.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
2. V. Sze, Y. Chen, J. Emer, A. Suleiman, and Z. Zhang, "Hardware for machine learning: challenges and opportunities," in *IEEE Custom Integrated Circuits Conference* (2017), pp. 1–8.
3. J. Goodman, *Introduction to Fourier Optics*, 3rd ed. (Roberts, 2005).
4. Y. Shen, N. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
5. S. Colburn, Y. Chu, E. Shilzerman, and A. Majumdar, "Optical frontend for a convolutional neural network," *Appl. Opt.* **58**, 3179–3186 (2019).
6. J. Feldmann, N. Youngblood, C. Wright, H. Bhaskaran, and W. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–214 (2019).
7. V. Bangari, B. Marquez, H. Miller, A. Tait, M. Nahmias, T. Lima, H. Peng, P. Prucnal, and B. Shastri, "Digital electronics and analog photonics for convolutional neural networks (DEAP-CNNs)," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7701213 (2020).
8. X. Lin, Y. Rivenson, N. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
9. J. Chang, V. Sitzmann, X. Dun, W. Heidrich, and G. Wetzstein, "Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification," *Sci. Rep.* **8**, 12324 (2018).
10. Y. Zuo, B. Li, Y. Zhao, Y. Jiang, Y. Chen, P. Chen, G. Jo, J. Liu, and S. Du, "All-optical neural network with nonlinear activation functions," *Optica* **6**, 1132–1137 (2019).
11. D. Ballarini, A. Gianfrate, R. Panico, A. Opala, S. Ghosh, L. Dominici, V. Ardizzone, G. Milena, G. Lerario, G. Gigli, T. Liw, M. Matuszewski, and D. Sanvitto, "Polaritonic neuromorphic computing outperforms linear classifiers," *Nano Lett.* **20**, 3506–3512 (2020).
12. A. Ryou, S. Colburn, and A. Majumdar, "Image enhancement in a miniature self-imaging degenerate optical cavity," *Phys. Rev. A* **101**, 013824 (2020).
13. S. Colburn, A. Zhan, and A. Majumdar, "Varifocal zoom imaging with large area focal length adjustable metalenses," *Optica* **5**, 825–831 (2018).
14. K. Matsushima and T. Shimobaba, "Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields," *Opt. Express* **17**, 19662–19673 (2009).
15. G. Agrawal, *Nonlinear Fiber Optics*, 5th ed. (Academic, 2013).
16. J. Davis, D. Cottrell, J. Campos, M. Yzuel, and I. Moreno, "Encoding amplitude information onto phase-only filters," *Appl. Opt.* **38**, 5004–5013 (1999).
17. H. Noh, J. Mun, and B. Han, "Regularizing deep neural networks by noise: its interpretation and optimization," in *Advances in Neural Information Processing Systems* (2017), pp. 5109–5118.
18. N. Nagabushan, N. Satish, and S. Raghuram, "Effect of injected noise in deep neural networks," in *IEEE International Conference on Computational Intelligence and Computing Research (ICCI)* (2016), pp. 1–5.
19. B. Poole, J. Sohl-Dickstein, and S. Ganguli, "Analyzing noise in autoencoders and deep networks," arXiv:1406.1831 (2014).
20. L. Li, P. Patki, Y. Kwon, V. Stelmakh, B. Campbell, M. Annamalai, T. Lakoba, and M. Vasilyev, "All-optical regenerator of multi-channel signals," *Nat. Commun.* **8**, 884 (2017).
21. O. Tzang, E. Niv, S. Singh, S. Labouesse, G. Myatt, and R. Piestun, "Wavefront shaping in complex media with a 350 kHz modulator via a 1D-to-2D transform," *Nat. Photonics* **13**, 788–793 (2019).
22. G. Gallego, T. Delbruck, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: a survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, <https://ieeexplore.ieee.org/document/9138762> (2020).
23. C. Foot, *Atomic Physics* (Oxford University, 2005).

# PHOTONICS Research

## Interfacing photonics with artificial intelligence: an innovative design strategy for photonic structures and devices based on artificial neural networks

YIHUAO XU,<sup>1</sup> XIANZHE ZHANG,<sup>2</sup> YUN FU,<sup>2,3</sup> AND YONGMIN LIU<sup>1,2,\*</sup> 

<sup>1</sup>Department of Mechanical and Industrial Engineering, Northeastern University, Boston, Massachusetts 02115, USA

<sup>2</sup>Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts 02115, USA

<sup>3</sup>Khoury College of Computer Science, Northeastern University, Boston, Massachusetts 02115, USA

\*Corresponding author: [y.liu@northeastern.edu](mailto:y.liu@northeastern.edu)

Received 15 December 2020; revised 31 January 2021; accepted 3 February 2021; posted 5 February 2021 (Doc. ID 417693); published 31 March 2021

Over the past decades, photonics has transformed many areas in both fundamental research and practical applications. In particular, we can manipulate light in a desired and prescribed manner by rationally designed subwavelength structures. However, constructing complex photonic structures and devices is still a time-consuming process, even for experienced researchers. As a subset of artificial intelligence, artificial neural networks serve as one potential solution to bypass the complicated design process, enabling us to directly predict the optical responses of photonic structures or perform the inverse design with high efficiency and accuracy. In this review, we will introduce several commonly used neural networks and highlight their applications in the design process of various optical structures and devices, particularly those in recent experimental works. We will also comment on the future directions to inspire researchers from different disciplines to collectively advance this emerging research field.

© 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.417693>

### 1. INTRODUCTION

Novel optical devices consisting of elaborately designed structures have become an extremely dynamic and fruitful research area because of their capability of manipulating light flow down to the nanoscale. Thanks to the advanced numerical simulation, fabrication, and characterization techniques, people are able to design, fabricate, and demonstrate dielectric and metallic micro- and nano-structures with sophisticated geometries and arrangements. For instance, metamaterials and metasurface comprising subwavelength structures, called meta-atoms, can show extraordinary properties beyond those of natural materials [1]. Many metadevices have been reported that offer enormous opportunities for technology breakthroughs in a wide range of applications including light steering [2–5], holography [6–9], imaging [10–14], sensing [15–17], and polarization control [18–21].

At present, we can handle most of the photonic design problems by accurately solving Maxwell's equations using numerical algorithms such as the finite element method (FEM) and finite-difference time-domain (FDTD) method. However, those methods often require plenty of time and computational resources, especially when it comes to the inverse design problem

aiming to retrieve the optimal structure from target optical responses and functionalities. In the conventional procedure, we normally start with full-wave simulations of an initial design based on the empirical knowledge and then adjust the geometric/material parameters iteratively to approach the customer-specific requirements. Such a trial-and-error process is time consuming, even for most experienced researchers. The initial design strongly relies on our experience and cognition, and usually some basic structures are chosen, including splitting resonators [22,23], helix [24], cross [25], bowtie [26], L-shape [2], and H-shape [27,28] structures. Although it is known that a specific type of structures can produce a certain optical response (e.g., strong magnetic resonance from splitting resonators and chiroptical response from helical structures), sometimes the well-established knowledge may limit our aspiration to seek an entirely new design that is suitable for the same applications or even more complicated ones when the traditional approach is not applicable.

Artificial neural networks (ANNs) provide a new and powerful approach for photonic designs [29–37]. ANNs can build an implicit relationship between the input (i.e., geometric/materials parameters) and the output (i.e., optical responses),

mimicking the nonlinear nerve conduction process in the human body. With the help of well-trained ANNs, we can bypass the complicated and time-consuming design process that heavily relies on numerical simulations and optimization. The functions of most ANN models for photonic designs are two-fold: the forward prediction and inverse design. The forward prediction network is used to determine the optical responses from the geometric/material parameters, and it can serve as a substitute for full-wave simulations. The inverse design network aims to efficiently retrieve the optimal structure from given optical responses, which is usually more important and challenging in the design process. One main advantage of the ANN models is the speed. For example, producing the spectrum of a meta-atom from a well-trained forward prediction model only takes a few milliseconds, orders of magnitude faster than typical full-wave simulations based on FEM or FDTD [38–40]. In the meantime, the accuracy of the ANN models is comparable with rigorous simulations. For instance, the mean squared loss of spectrum prediction is typically on the order of  $10^{-3}$  to  $10^{-5}$  [40,41]. Moreover, ANNs can unlock the nonintuitive and nonunique relationship between the physical structure and the optical response, and hence potentially enlighten the researchers with an entirely new class of structures.

Solving the photonic design problem by ANNs is a data-driven approach, which means a large amount of training sets with both geometric/material parameters and optical responses are needed. Once the ANN model works well on the training data set, it can be tested on a test set or real problem. The test and training data sets should be in the same design framework but contain completely different data. The general workflow for a forward prediction network includes four steps. First, a large number of input structures and output optical responses are generated from either simulations or experiments. In most of the published works, the amount of data is on the order of  $10^4$ . It is noted that the performance of the neural networks depends on both the size and quality of data. To improve the quality of training data, some researchers have applied rule-based optimization methods in the generation of initial training data [42] or attempted to progressively increase the dimension of the training data with the new ones from the trained model [43]. Then we design the ANNs with a certain network structure, such as fully connected layers (FCLs)-based neural networks or convolutional neural networks (CNNs). Next, the training data set is fed into the network, and we optimize the weight and bias for each node. Finally, the well-trained ANNs can be used to predict the response of other input structures that are outside the training and test data sets. As for the inverse design problem, one can simply reverse the input and output and use a similar network structure. However, for some problems, it requires complex methods and algorithms.

This review is devoted to the topic of designing photonic structures and devices with ANNs. We will focus on very recent works on this topic, especially the experimental demonstrations, after introducing the widely used ANNs. The remaining part of the review is organized as follows. In Section 2, we will discuss the basic FCLs and their application in the prediction of design parameters. Then, in Section 3, we will focus on the CNNs that are used in the retrieval of much more complicated

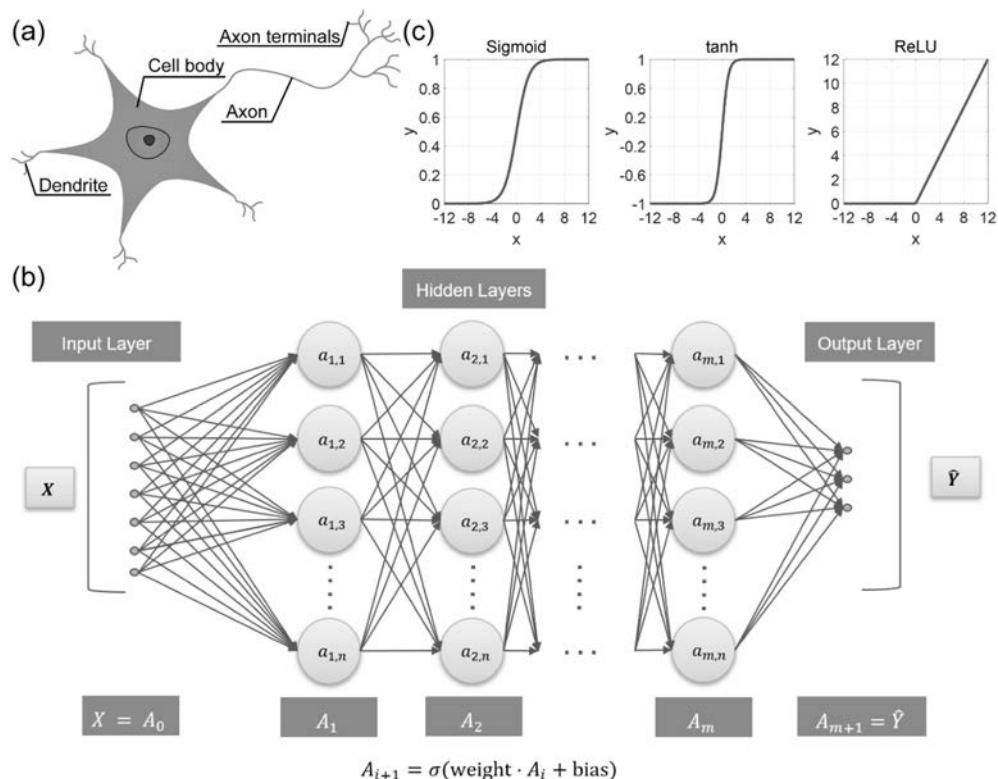
structures described by pixelated images. In Section 4, other useful and efficient hybrid algorithms by combining deep learning and conventional optimization methods for photonic design will be discussed. In the last section, we will conclude the review by discussing the achievements, current challenges, and outlooks in the future.

## 2. PHOTONIC DESIGN BY FULLY CONNECTED NEURAL NETWORK

### A. Introduction of FCLs

In the nervous system, the electric signals and information are transited by neurons. Figure 1(a) is the schematic illustration of a neuron, in which the main components include the dendrite, cell body, and axon. The dendrites receive and integrate signals from other neurons. Once the signal is strong enough, the cell is activated and then passes the signal to the next neuron through the axon. In analogy to biological neurons, spiking neural networks (SNNs) were introduced decades ago [44,45]. In SNNs, not all neurons are activated at each propagation loop. Only when the action potential, which mimics the membrane potential, reaches a certain value will the neurons transmit information to the next neuron. The FCLs, also called dense layers, are mathematically simplified structures in comparison with the SNNs, in which any neurons in adjacent layers are connected as shown in Fig. 1(b). The FCLs-based network consists of an input layer, hidden layers, and an output layer. Each layer receives the input from the preceding layer and combines the signal with a tensor operation that takes the weights and biases as learnable parameters. This is a purely linear process. The combination of linear calculation of all hidden layers is still linear, and it is not able to build complex relationships between the input and output. In this context, a nonlinear “activation” process is essential for the neurons. Therefore, the activation function  $\sigma(\cdot)$  should be a nonlinear function and properly selected so that the ANNs can approximate arbitrary functions. The Sigmoid, tanh and ReLU functions plotted in Fig. 1(c) are three commonly used activation functions. The Sigmoid and tanh functions map the input into (0,1) range. When the input is extremely small (or extremely large), the gradient of the two functions vanishes, which is not favorable for the ANNs. The ReLU is widely used to overcome this issue in the positive range, and it is not computationally expensive compared to Sigmoid and tanh functions.

The training process of the fully connected neural network is quite straightforward. The training set contains an input vector  $X$  and an output vector  $Y$  ( $Y$  can be a vector of complex/real values for regression problems or vector of discrete integers as labels for classification problems). The performance of the model is highly dependent on the quantity and quality of the training data set. During the training process, the network first takes the vector  $X$  as input and calculates the output  $\hat{Y}$  through the tensor operation and activation from left to right. Then a loss function (or cost function) is defined and needs to be minimized in order to calculate the performance of the neural network. For instance, we can use mean-squared-error (MSE) [ $\text{loss}(Y, \hat{Y}) = (Y - \hat{Y})^2$ ] for regression problems and cross-entropy loss [ $\text{loss}(Y, \hat{Y}) = -Y^T \cdot \log(\hat{Y})$ ] for classification problems. The next step, the backpropagation of error,



**Fig. 1.** (a) Illustration of a biological neuron. (b) FCLs-based neural network, in which all neurons in adjacent layers are connected. (c) Three widely used activation functions: Sigmoid, tanh, and ReLU.

is the most critical part of ANNs. In the ANN, there are a series of learnable parameters to be optimized, i.e., the weight and bias of each layer. We can then derive the partial derivative of the loss with respect to each parameter  $\frac{\partial \text{loss}(Y, \hat{Y})}{\partial \text{weight}}$ ,  $\frac{\partial \text{loss}(Y, \hat{Y})}{\partial \text{bias}}$ . To calculate those values, we need to apply the chain rule layer by layer from the end of the ANN to the front. This is why the process is called “backpropagation.” Finally, all the parameters are optimized by the stochastic gradient descent method:

$$\begin{cases} \text{weight}' = \text{weight} - \text{lr} \cdot \frac{\partial \text{loss}(Y, \hat{Y})}{\partial \text{weight}} \\ \text{bias}' = \text{bias} - \text{lr} \cdot \frac{\partial \text{loss}(Y, \hat{Y})}{\partial \text{bias}} \end{cases}$$

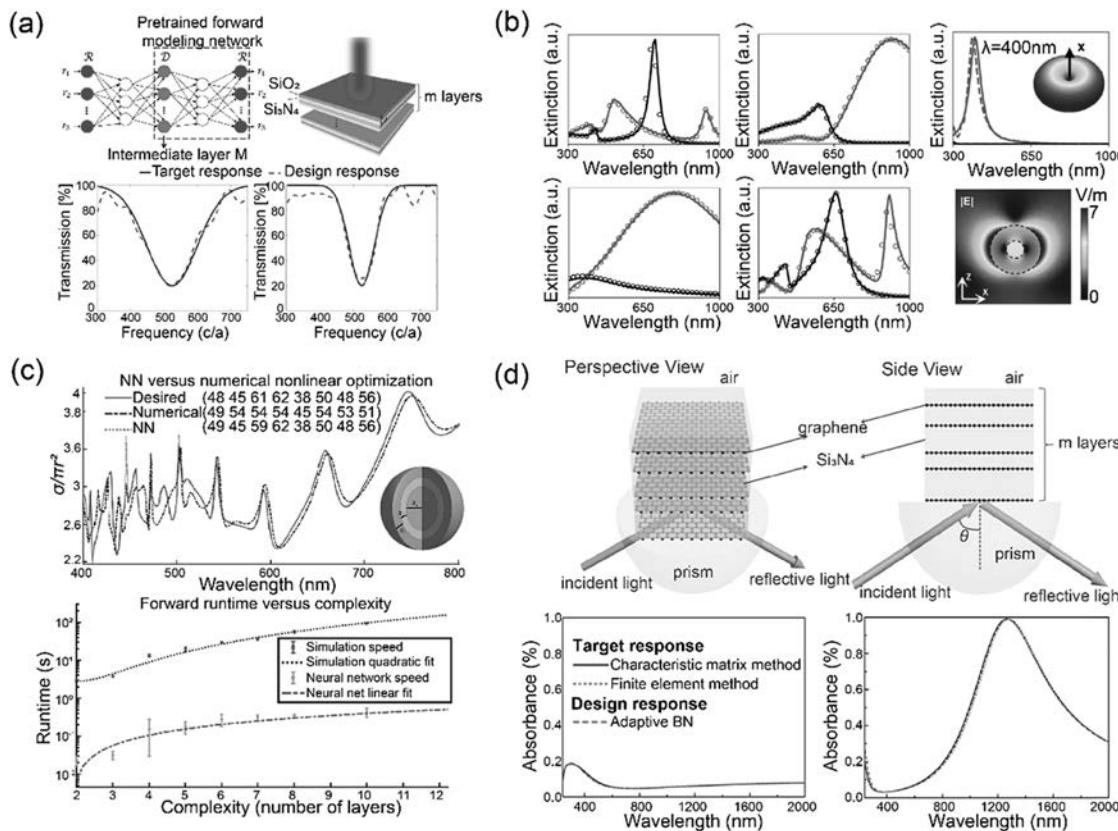
Here the learning rate lr is a hyperparameter that is usually controlled by the user and is not learnable. The training process is iterated until the loss is minimized. Different learning rates would result in different situations: a large learning rate will cause issues for the model to converge, while a small learning rate will increase the training time of the model. Therefore, the general approach is to assign a large learning rate at the beginning of the training, and after the model is trained for several epochs, the learning rate can be tuned to a smaller value.

## B. Design Parameterized Structure by FCLs-Based ANNs

FCLs have been extensively adopted to design optical devices, especially in the field of metasurface and nanostructure design. In early 2018, D. Liu *et al.* introduced, for the first time, a

tandem network architecture for the inverse design problem [46]. There is one fundamental challenge in training ANNs for inverse design, arising from the fact that very similar optical responses may be achieved by different structures. Such non-unique one-to-many mapping makes the neural network hard to converge if conflicting instances with almost the same optical responses but different geometric labels exist in the training data set. Mathematically, the gradient of the functions to be approximated by the ANNs is extremely large at this data point. To tackle this challenge, the authors proposed a network structure consisting of a pretrained forward model and inversed design FCLs, which is illustrated in the top panel of Fig. 2(a). The network structure avoids direct comparison between the retrieved geometric parameters. Instead, it compares the predicted spectra of the retrieved structures. Therefore, the prediction of the network will converge to only one structure that can satisfy the required spectra, solving the one-to-many problem in the inverse design. The authors used the tandem neural network to design dielectric multilayers composed of  $\text{SiO}_2$  and  $\text{Si}_3\text{O}_4$ . The results are plotted in the bottom panel of Fig. 2(a), in which the transmission spectra of the retrieved structure (green dashed line) can well match the desired Gaussian-shaped spectra (blue solid lines).

Subsequent works have further confirmed the good performance of the tandem network architecture. For instance, S. So *et al.* used a similar ANN structure to design core-shell structures (with three layers) that support strong electric and magnetic dipole resonances [47]. The ANN was built to learn



**Fig. 2.** (a) Top: Schematic of the tandem neural network and  $\text{SiO}_2$  and  $\text{Si}_3\text{N}_4$  multilayers. Bottom: Two examples of target spectra (blue solid lines) and simulated spectra of retrieved structures (green dashed lines). The target spectra are in a Gaussian shape. (b) Left: Predicted (open circles) extinction cross section of the electric dipole (red) and magnetic dipole (black) of core-shell nanoparticles. The solid lines are target responses. Right: Simulated extinction spectra and the corresponding electric field distribution of core-shell nanoparticles. (c) Top: Simulation result and inverse design prediction of the scattering cross section of core-shell nanoparticles. Bottom: Runtime comparison between the conventional method and neural network. (d) Top: A multilayer structure composed of  $\text{Si}_3\text{N}_4$  and graphene. Bottom: Optical response of the designed nanostructures (with either low/near-unity absorbance in graphene) under the excitation of s-polarized light. (a) is reproduced from Ref. [46] with permission; (b) is reproduced from Ref. [47] with permission; (c) is reproduced from Ref. [38] with permission; (d) is reproduced from Ref. [48] with permission.

the correlation between the extinction spectra and core-shell nanoparticle designs, including the material information and shell thicknesses. In Fig. 2(b), the predicted (open circles) extinction cross sections of the electric dipole (red) and magnetic dipole (black) of core-shell nanoparticles are compared with the target responses (solid lines). It is clear that both the electric dipole and magnetic dipole spectra of the designed core-shell nanoparticles fit well with the expectations. J. Peurifoy *et al.* also studied the inverse design with ANNs for multilayered particles (up to eight layers), with a focus on the scattering spectra [38]. The FCLs were used in both forward prediction of scattering cross-section spectra and the inverse design from the spectra. Using a model trained with 50,000 training data, they can achieve a mean relative error of around 1%. One example is shown in the top panel of Fig. 2(c), in which the result from the neural network is compared with numerical nonlinear optimization as well as the desired spectra. The comparison demonstrates that the neural network model performs better in this design problem. Moreover, the running time of the ANNs-aided inverse design is

shortened by more than 100 times in comparison with full-wave simulation as demonstrated in the bottom panel of Fig. 2(c). This result clearly shows the advantage of ANNs in terms of efficiency.

Besides the tandem network, other approaches have been introduced to improve the performance of the FCLs-based neural network. In 2019, Y. Chen *et al.* employed an adaptive batch-normalized (BN) neural network, targeting the smart and quick design of graphene-based metamaterials as illustrated in the top panel of Fig. 2(d) [48]. Specifically, a layer using an adaptive BN algorithm is placed before each hidden layer to overcome the limitation of BN in small sampling spaces. In the adaptive BN network, it takes activation  $h_i$  of each neuron in a minibatch  $B$ , batch normalization parameters  $\gamma$ ,  $\delta$ , and adaptive parameters  $\alpha$ ,  $\beta$  as the inputs. The outputs of the system are the new activation  $\hat{h}_i$  for each neuron. The authors tested their method by deriving the thickness of each  $\text{Si}_3\text{O}_4$  layer in the structures. Prediction accuracy of over 95% was achieved. The bottom panel of Fig. 2(d) plots the optical responses of two different examples with varied absorbance in

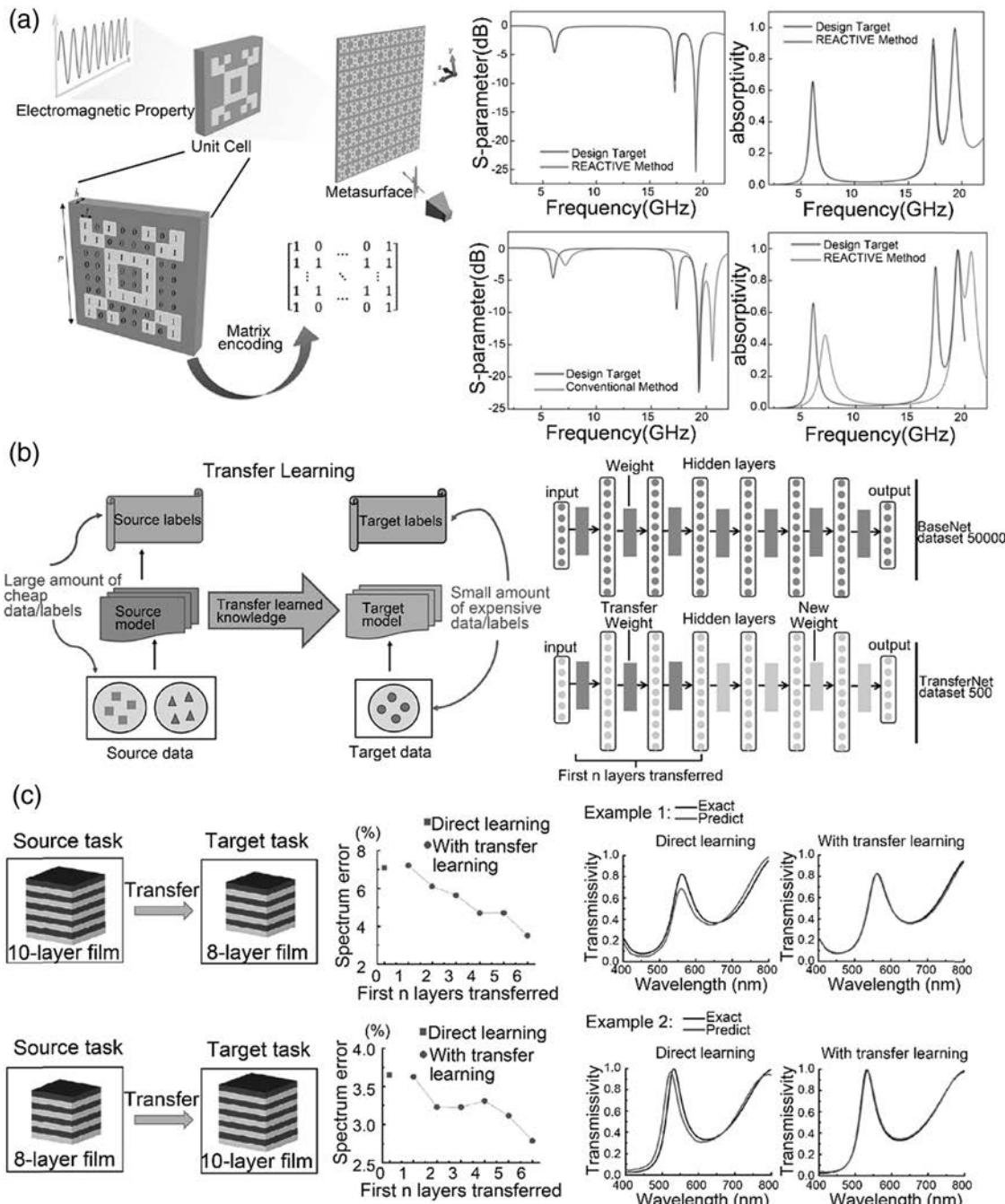
graphene, showing excellent accordance between the target and design responses.

In parallel, T. Qiu *et al.* proposed a new method, named REACTIVE, to conduct the inverse design based on reflection spectra [39]. The authors applied this method to inversely design a metasurface whose unit cell can be described as a matrix of  $8 \times 8$  as shown in the left panel of Fig. 3(a). The input data sets are preprocessed by Gaussian smoothing and then transformed by a discrete cosine transform that can be modeled as

$$F(u) = c(u) \sum_{i=0}^{N-1} f(i) \cos \left[ \frac{(2i+1)\pi}{2N} u \right],$$

$$c(u) = \begin{cases} \sqrt{1/N}, & u = 0 \\ \sqrt{2/N}, & u \neq 0 \end{cases}$$

In this model, the S-parameters of the desired structure are the needed output. Once the S-parameters are generated by the trained deep learning network, the matrix of the designed



**Fig. 3.** (a) Left: Schematic illustration of the metasurface, the unit cell, and matrix encoding method. Right: Predicted S-parameter and absorptivity with the REACTIVE method. (b) Illustration of the neural network architecture consisting of BaseNet and TransferNet. (c) The trend of spectrum error when  $n$  layers are transferred to the TransferNet and the predicted transmission spectra for two examples. (a) is reproduced from Ref. [39] with permission; (b) and (c) are reproduced from Ref. [49] with permission.

metasurface will be automatically generated by REACTIVE. In Fig. 3(a), the right panel shows the results from REACTIVE, including S-parameter  $S_{11}$  (i.e., reflection coefficient) and absorptivity, which perfectly match the design targets.

Due to the data-driven nature of deep learning, the performance of a well-trained ANN highly relies on the training set, and the prediction loss is likely to increase as the inputs deviate from the training set. Therefore, a challenge in the deep-learning-aided inverse design lies in extending the capability of ANNs to an alternated data set that is very different from the training data. Usually, one needs to generate an entirely new training set for similar but different physical scenarios. In this context, reducing the demand for computational data is an efficient way to accelerate the training of deep learning models. Y. Qu *et al.* proposed a transfer learning method, which is schematically illustrated in Fig. 3(b), to migrate knowledge under different physical scenarios [49]. The prediction accuracy is significantly improved, even with a much smaller data set for new tasks. Two sets of ANNs are involved in this work. The first one, named BaseNet, is trained with initial data. The second one, called TransferNet, copies the first  $n$  layers from the BaseNet, and the entire system is fine-tuned simultaneously. The authors first transferred the spectra prediction task from a 10-layer film to an 8-layer film, where the source and target task were trained with 50,000 and 5000 examples, respectively. Comparing to direct learning, the result is good enough since the error drops when  $n$  increases as shown in Fig. 3(c). The TransferNet is applicable for different structures, ranging from multilayer nanoparticles to multilayer films. Based on the model, a multitask learning scheme was studied, which combined the learning for multiple tasks at the same time. It was shown that the neural network in conjunction with the transfer learning method can produce more accurate predictions.

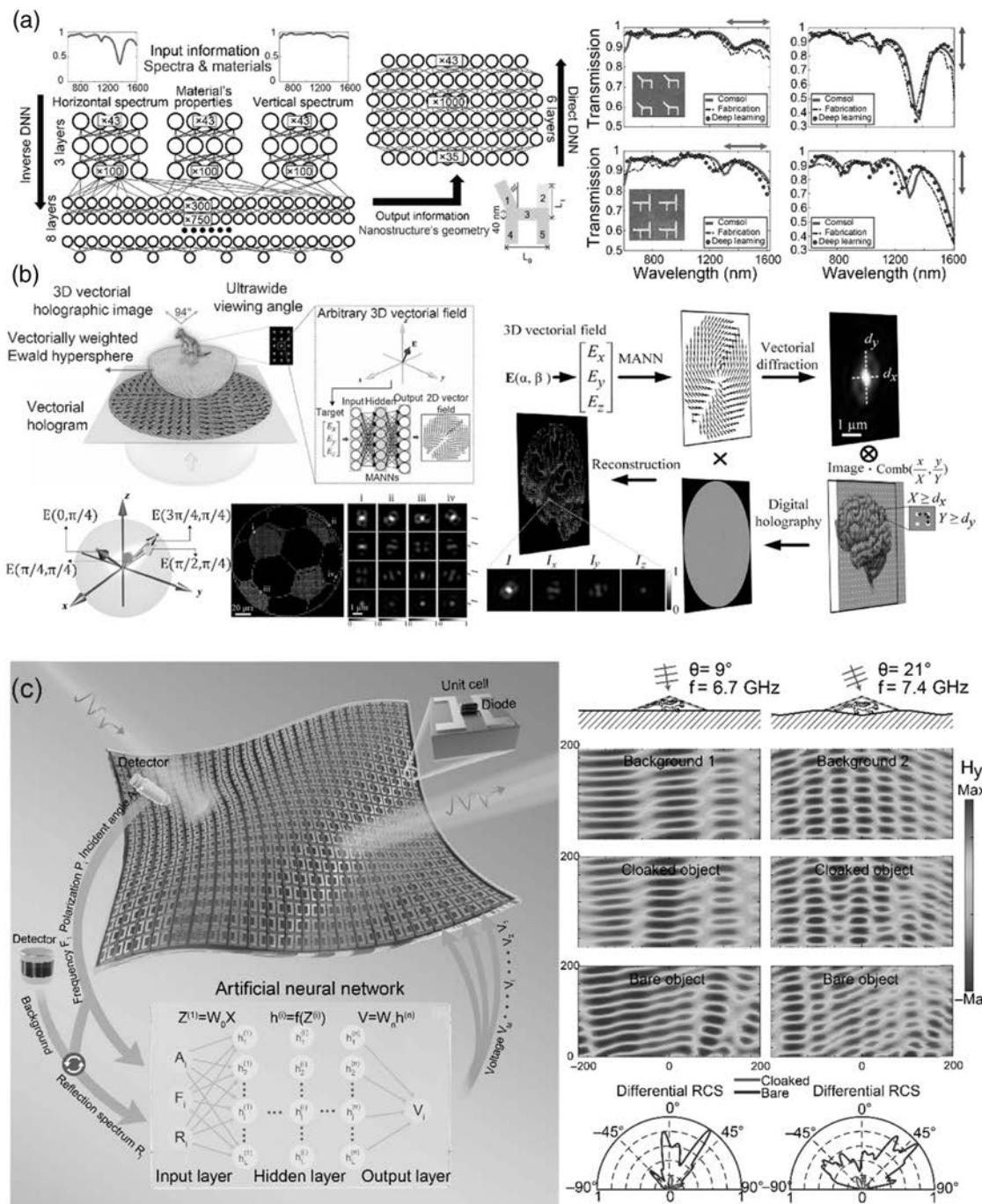
The FCLs have also been utilized in reinforcement learning [50–53], which is another hot area of machine learning, for the inverse design problem. Reinforcement learning has already achieved great performance in robotics, system control, and game-playing (AlphaGo). Instead of predicting the optimized geometry, the ANNs in reinforcement learning behave as an iterative optimization method. In each step, an action to optimize the geometry parameters is predicted. For instance, the action can be increasing or decreasing several parameters by a certain value. The advantage of this approach is that it can be adaptive to specific problems, and it can provide guidance for conventional trial-and-error optimization methods.

People have devoted experimental efforts in conjunction with the development of general models and algorithms using ANNs. For example, I. Malkiel *et al.* experimentally demonstrated that a deep neural network trained with thousands of synthetic experiments can retrieve subwavelength meta-atoms from far-field measurements and address the inverse design problem [54]. In their work, the training of the inverse network to predict the structure based on the transmission spectra served as the first step. The material properties were also considered additional inputs. The second step was to train the direct network for forward prediction on top of the first

network as shown in the left panel of Fig. 4(a). A significant and encouraging improvement in accuracy was noted when using eight joint layers. Based on the far-field spectra and the developed neural networks, the authors were able to derive the geometries of nanostructures. They achieved great agreement between the desired spectra and simulated spectra of the fabricated samples as shown in the right panel of Fig. 4(a).

In addition to spectrum prediction [55,56], the FCLs-based ANNs have also been used in the inverse design to realize other functionalities and benefit real-world applications [57–62]. Holographic images, for example, can be optimized by ANNs to achieve a wide viewing angle and three-dimensional vectorial field as recently demonstrated by H. Ren *et al.* [63]. They used a network named multilayer perceptron ANN (MANN), which was composed of an input layer fed with an arbitrary three-dimensional (3D) vectorial field, four hidden layers, and an output layer for the synthesis of a two-dimensional (2D) vector field. There are 1000 neurons within each hidden layer. The scheme of this ANN is shown in the top left panel of Fig. 4(b). The authors showed that an arbitrary 3D vectorial field can be achieved with a 2D vector field predicted by the well-trained model. A 2D Dirac comb function was then applied to sample the desired image. Subsequently, digital holography, calculated from the desired image, was combined with the 2D vector field. This process can be visualized in the right panel of Fig. 4(b). With a split-screen spatial light modulator that independently controls the amplitude and phase orthogonal circularly polarized light, any desired 2D vector beam can be generated. As a result, the experimentally measured image from the hologram can show four different 3D vectorial fields in different regions as presented in the bottom left panel of Fig. 4(b). The authors experimentally realized an ultrawide viewing angle of 94° and high diffraction efficiency of 78%. The demonstrated 3D vectorial holography opens avenues to widespread applications such as holographic display as well as multidimensional data storage, machine learning microscopy, and imaging systems.

Another exciting work enabled by ANNs is a self-adaptive cloak that can respond within milliseconds to ever-changing incident waves and surrounding environments without human intervention [64]. A pretrained ANN was adopted to achieve the function. As schematically illustrated on the left panel of Fig. 4(c), at the surface of the cloak, a single layer of active meta-atoms was applied, and the reflection spectrum of each varactor diode was controlled by DC bias voltage independently. To achieve the invisibility cloak function, the bias voltage was determined by the pretrained ANN with the incident wave characteristics (such as the incident angle, frequency, and reflection amplitude) as the input. The temporal response of the cloak was simulated, and an extremely fast transient response of 16 ms can be observed in the simulation. The authors then conducted the experiment, where a p-polarized Gaussian beam illuminated at an angle  $\theta$  on a chameleon object covered by the cloak. Two detectors were used to extract the signals from the background and the incident wave to characterize the cloak. The right panel of Fig. 4(c) shows the experimental results at two incident angles (9° and 21°) and two frequencies



**Fig. 4.** (a) Left: Architecture of the proposed neural network for nonlinear layers. Right: Predicted, simulated, and measured transmission spectra of two gold nanostructures under different polarization conditions. (b) Left: Illustrations of MANN used for reconstruction of 3D vectorial field. Right: Experimental approach and characterizations of 3D vectorial holography based on a vectorial hologram. (c) Left: Schematic of a deep-learning-enabled self-adaptive metasurface cloak. Right: Demonstration of the self-adaptive cloak response subject to random backgrounds and incidence with varied angles and frequencies. (a) is reproduced from Ref. [54] with permission; (b) is reproduced from Ref. [63] with permission; (c) is reproduced from Ref. [64] with permission.

(6.7 and 7.4 GHz). The magnetic field distribution in the case of a cloaked object is similar to that when only the background is present, while it is distinctly different from the bare object case. Differential radar cross-section (RCS) measurement further confirmed the performance of the cloak.

### 3. RETRIEVE COMPLEX STRUCTURES BY CONVOLUTIONAL NEURAL NETWORKS

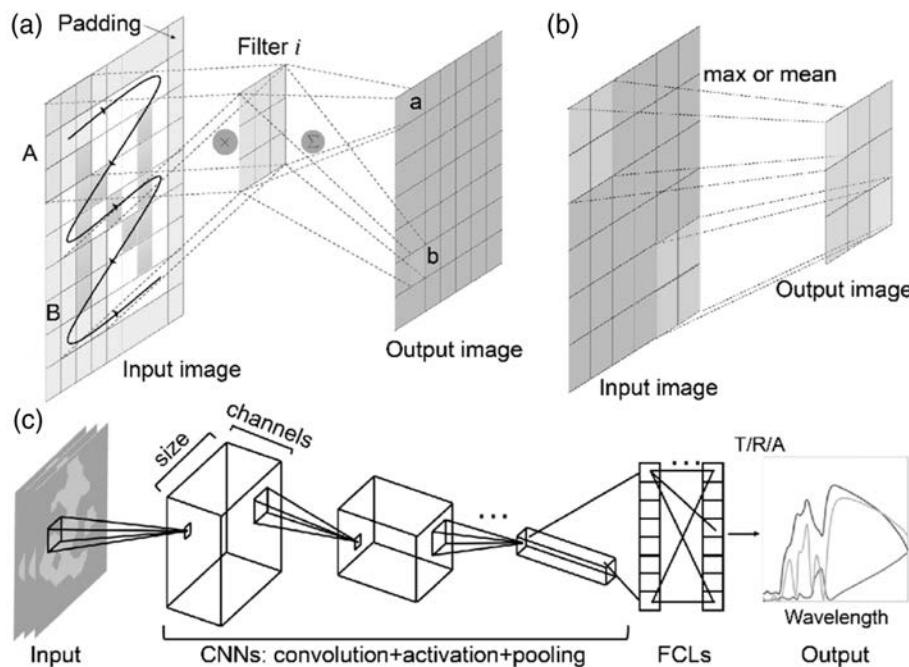
#### A. Introduction of CNNs

The desired designs and structures are oftentimes hard to parameterize, especially when the structure of interest contains

many basic shapes [41,65] or is freeform [66,67]. In some cases, we need to deal with complex optical responses as the input [68]. Therefore, converting the structure to a 2D or 3D image is usually a good approach in these studies. Moreover, it can offer much larger degrees of freedom in the design process. However, preprocessing is required to handle the image input if we still want to use the FCLs-based model. Reshaping the image to a one-dimensional vector and applying feature extraction with linear embeddings, such as principal component analysis and random projection, are two effective ways to preprocess the image so that the input is compatible with the FCLs. However, the performance is usually not satisfactory. The reason is that these conversions will either break down the correlation of the nearest pixels in the vertical direction within an individual image or miss part of the information describing the integrality of the whole image. An extremely large dimension of the input is another big issue, which will increase the number of connections between layers quadratically. For conventional parameter input, the input dimension is usually a few tens or hundreds, while for a vectorized image, even an image with  $64 \times 64$  pixels will result in a 4096-dimensional input vector. CNNs are very suitable to deal with such circumstances. CNNs accept an image input without preprocessing, and then several filters move along the horizontal and vertical directions of the image to extract different features. Each filter has a certain weight to perform a convolutional operation at each subarea of the image, that is, the summation of the pointwise multiplication between the value of the subarea and the weight of the filter.

To explain the function of CNNs in detail, let us assume an input  $(C, X, Y)$ . Here  $C$  is the number of channels of an image, while  $X$  and  $Y$  are the number of pixels in horizontal and

vertical directions, respectively. For binary or gray images  $C = 1$ , and for RGB images  $C = 3$ . Then each CNN consists of a weight tensor that has  $N_f$  filters with the dimension  $(C, X_f, Y_f)$ , meaning each filter is built with  $C$  channels of an  $X_f \times Y_f$  matrix (usually a  $3 \times 3$  or  $5 \times 5$  matrix is used). The CNN is normally built with three operations, including convolution, activation, and pooling (sometimes a batch normalization will be added). Figure 5(a) illustrates the convolution operation (consider  $C = 1$ ). Each filter is initially placed on the top left  $X_f \times Y_f$  subarea of each image. The pointwise multiplication of the two  $X_f \times Y_f$  matrices is calculated and summed to a single value in the output image. Then the filter moves a certain number of pixels (known as “stride”) and repeats the process until the whole image is mapped to the output. The dimension of the output is usually smaller than that of the input. However, the output dimension can be easily tuned by adding paddings to the input images, which expand the dimension of the input image with zero pixels. In this example where one round of padding is added, the output image will have the same dimension as the input (stride equals 1, and the filter dimension is  $3 \times 3$ ). The activation function plays a significant role in the CNNs for the same reason as FCLs, and we can choose similar functions as previously mentioned. A pooling layer helps to reduce the dimension of the image. It usually maps a  $2 \times 2$  (or  $3 \times 3$ ) area in the input to a single value in the output according to the maximum or mean value of the four (or nine) values, as represented in Fig. 5(b). The entire workflow for conventional CNNs is shown in Fig. 5(c). The inputs are several images, and each represents a certain design of structure. The inputs pass through layers of CNN with three operations, and the size of the tensor gradually shrinks while the



**Fig. 5.** (a) Schematic of the convolution operation, in which the filters map the subarea in the input image to a single value in the output image. (b) Schematic of the pooling operation, in which the subarea in the input image is pooled into a single value in the output according to the maximum or mean value. (c) The workflow of a conventional CNN. The input images pass through several CNNs, and then the extracted features are passed into the FCLs to predict the response (e.g., transmission, reflection, and absorption spectra).

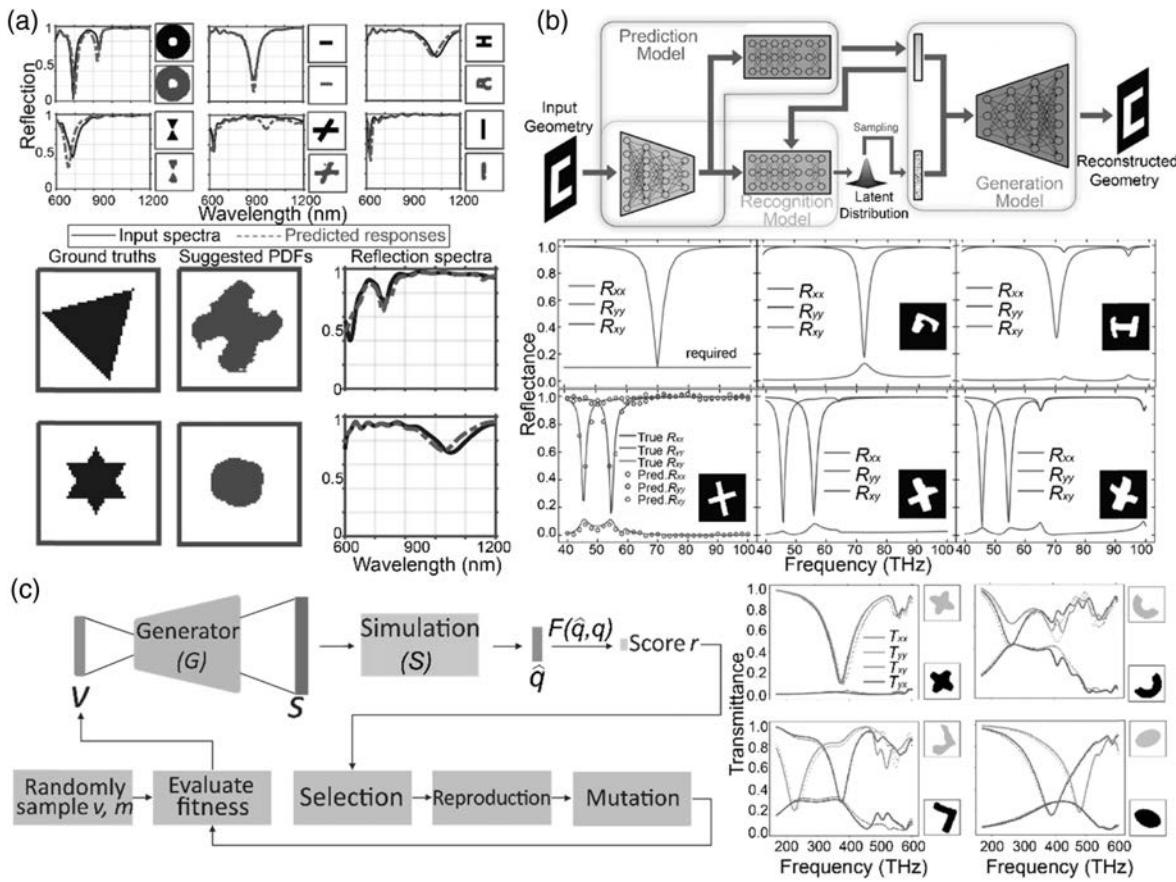
number of channels expands. The output now becomes a 1D vector. It can be regarded as the features extracted from the image, and the features are fed into the FCLs to predict the final output that is related to the optical response. The MSE and cross-entropy loss discussed in the previous section can also serve as the loss function in many cases of CNNs. The loss calculated by comparing the predicted and true response undergoes backpropagation through all layers to update the parameters. We want to emphasize that other loss functions, such as Kullback–Leibler divergence [41] and mean absolute error [69], can also be used in ANNs, depending on the physical constraints and the expected functions of the ANNs.

### B. Design Complex Photonic Structures by CNNs

The CNNs have greatly expanded the design space of the possible structures that one can explore. For example, plasmonic structures have been extensively studied over the past decades, due to their unique features in optics and photonics and far-reaching impacts on other disciplines [70–74]. By carefully designing the geometry and composite materials, we can confine light into a sub-10 nm dimension with the local field amplified by 10–1000 times at the resonant wavelengths. Therefore, building a relationship between the design of the plasmonic structure and the corresponding optical responses is of great interest. In the work of I. Sajedian *et al.* published in 2019, the authors combined the CNNs with the recurrent neural networks (RNNs) to predict the absorption spectra of complex plasmonic structures in the near-infrared region [40]. The CNNs helped to extract the features from the pixelated structures, and the RNNs with gated recurrent unit layers were used to predict the spectra. The model showed an MSE loss lower than  $10^{-4}$  when training with 100,000 data. The authors have also examined the output after each layer to investigate how the higher-level features could be extracted as the model goes deeper. In the same year, S. So *et al.* reported the use of conditional deep convolutional generative adversarial networks (cDCGANs) to retrieve silver plasmonic structures with six basic shapes, such as circle, square, and cross, from given reflection spectra under linearly polarized illumination [69]. The generative adversarial networks (GANs) consist of a generator network and a discriminator network [65,75]. The training process for the GANs-based model is a competition between the generator and the discriminator. The generator generates structures from the input spectrum and a noise vector, trying to fool the discriminator that the generated structure is a rational structure according to the knowledge learned from the training set. The noise vectors are sampled from a conditional distribution, which is dependent on the prescribed spectra in this case. The discriminator tries to discriminate the “fake” structures generated from the generator and the “true” structures in the training data set. In the beginning, each input structure is pixelated into a  $64 \times 64$  image, and the CNNs are used to extract the features of the images in both networks. After running several epochs of the training process, even the optimized discriminator can hardly distinguish the difference between “fake” and “true” inputs, since the generator can generate extremely similar structures to the desired ones, resulting in a good model for inverse design. As shown in the top panel of Fig. 6(a), the simulated spectra of the retrieved

structures (red line) agree well with the desired spectra (black line), which are either simulated with an existing structure (first row) or randomly generated with a Lorentzian shape (second row). The overall accuracy is noticeable, reaching a 0.0322 mean-absolute error among 12 test samples after the model is trained with 10,150 training data. The authors also showed that the model can inversely design different structures (but are still within the basic shape groups), while the spectra meet the target as illustrated at the bottom of Fig. 6(a). The emergence of structures different from the ground truths can be attributed to the one-to-many mapping issue that we have discussed in the introduction section.

W. Ma *et al.* also demonstrated a probabilistic approach for the inverse design of plasmonic structures in 2019 [41]. In this work, the structure of interest was a metal-insulator-metal (MIM) structure, with geometries pixelated into  $64 \times 64$  images as training data. The authors focused on the co- and cross-polarized reflection spectra in the mid-infrared region from 40 to 100 THz. The developed neural network is shown at the top of Fig. 6(b), which comprises the prediction, recognition, and generation models. Again, the input geometry passes through the CNNs to extract the features from the image. Then the prediction model with FCLs can automatically predict the reflection spectra from the geometry features. For the inverse design part, the authors incorporated a variational auto-encoder (VAE) structure [76,77], which is a probabilistic approach, in the model. It works in the following way. First, the recognition network encodes both the structures and corresponding spectra into a latent space with a standard Gaussian prior distribution. While in the generation model, the network takes the desired spectra together with a latent variable randomly sampled from the conditional latent distribution to reconstruct one geometry. Here, the three models are trained together in an end-to-end manner. The well-trained model can not only predict the spectra from the given structure, serving as a powerful alternative for numerical simulation, but also reconstruct multiple structures from user-defined spectra. The bottom part of Fig. 6(b) shows the performance of the model trained with 30,000 data for spectral prediction and the inverse design for both user-defined spectra (first row) and spectra from a test structure (second row). The first column in the figure shows the target spectra. In the case where a test structure is used to generate the spectra, the predicted spectrum from the prediction model is also plotted as a scatter plot, which shows great coincidence with the spectra from full-wave simulation (solid lines). In the second and third columns, two examples of the geometry from the inverse design model and their simulated spectra are depicted. One can find that even though the structures are very different from each other and also from the ground truth, the spectra resemble the target ones. The authors further expanded the basic shapes by transfer learning to enable the reconstruction of a wide range of geometry groups. The generality of the model was exemplified by the designs of double-layer chiral metamaterials. Very recently, W. Ma and Y. Liu developed a semi-supervised learning strategy to accelerate the training data generation process, the most time-consuming part of the deep-learning-aided inverse design [78]. In addition to the labeled data that have both the geometries of



**Fig. 6.** (a) Top: Examples of cDCGAN-suggested images and the simulation results. Bottom: Entirely new structures suggested by the cDCGAN for desired spectra. (b) Top: The proposed deep generative model for metamaterial design, which consists of the prediction, recognition, and generation models. Bottom: Evaluation of the proposed model. The desired spectra either generated with user-defined function or simulated from an existing structure are plotted in the first column. The reconstructed structures with the simulated spectra are plotted in the second and third columns. (c) Left: Flowchart of the VAE-ES framework. Right: Test results of designed photonic structures from the proposed model and the simulated spectra. (a) is reproduced from Ref. [69] with permission; (b) is reproduced from Ref. [41] with permission; (c) is reproduced from Ref. [79] with permission.

structures and simulated spectra, the unlabeled data with only the geometry information are included. Unlike the labeled data where simulated spectra can be the input in the inverse design model, the predicted spectra of the unlabeled data are used as input to reconstruct the geometry. Without numerical simulation, the unlabeled data can be generated several orders of magnitude faster. They also help to dramatically lower the training loss by 10%–30% for the model trained with the same number of labeled data.

Z. Liu *et al.* introduced a hybrid approach by combining the VAE model and the evolution strategy (ES) [79]. The framework of the hybrid model is shown on the left of Fig. 6(c). In each iteration, a generation of latent vectors  $v$  is fed into the model and a structure is reconstructed. Then a well-trained simulator is used to predict the transmittance spectra of the structures, and the fitness score is calculated. If the criteria are not satisfied yet, the ES will perform reproduction and mutation with the mutation strength  $m$  to create a new generation of the latent vectors. Such a process is repeated until the criteria are met. The details of ES will be discussed in the genetic

algorithm part in the next section. The right panel of Fig. 6(c) shows the performance of the inverse design model. The solid line and dashed line are the simulated spectra of the test pattern (orange) by finite element method and the reconstructed pattern (black) from the hybrid model, respectively. All the works in Fig. 6 solve the one-to-many mapping issue with a probabilistic approach like VAEs and GANs, where a randomly sampled parameter or vector is combined with the desired optical response as the input to reconstruct the structure. It enables the ANNs to explore the full physical possibility of the design space to produce sophisticated structures for novel functions.

In 2019, Q. Zhang *et al.* demonstrated the digital coding metasurface using CNNs [80]. They explored different meta-atoms, each in the size of 8 mm and with  $16 \times 16$  pixels, to control the reflection phase. The CNNs model was built upon residual learning blocks and 70,000 training patterns. After training, the model can precisely predict the reflection phase; 90.05% of the test samples exhibited a deviation of less than  $2^\circ$  in the  $360^\circ$  phase range. Subsequently, the model was

used for the inverse design of meta-atoms with a prescribed phase response. More specifically, the goal was to create a 1-bit coding with two meta-atoms such that the reflection phase of the incident  $x$ - ( $p_x$ ) and  $y$ - ( $p_y$ ) polarized light satisfy

$$p_{xi} - p_{yi} = \theta, \quad |p_{y2} - p_{y1}| = 180^\circ.$$

This means that the two meta-atoms should have the same reflection phase difference  $\theta$  between cross-polarizations, and the relative phase between the two meta-atoms is maximized to  $180^\circ$ . With varied phase difference  $\theta$ , eight different 1-bit coding elements (with two structures for each coding) were predicted with  $45^\circ$  step in  $\theta$ . One example of the eight elements is plotted on the left of Fig. 7(a). By carefully combining the phase profile on a metasurface consisting of 16 designed units, the authors demonstrated the independent manipulation of the phase for orthogonal polarizations. As one example of potential applications, the authors fabricated several dual- and triple-beam coding metasurfaces that can deflect light with different polarizations into different angles at 10 GHz. The measurement was performed in the microwave chamber with a horn antenna as the excitation source. On the right of Fig. 7(a), we can find excellent agreement between the measured far-field scattering patterns and the simulated ones.

CNNs are widely applied in 2D image processing. The significance of CNNs is attributed to their ability to keep the local segment of the input as a whole, which can theoretically work in an arbitrary dimension. Taking advantage of this property, P. R. Wiecha and O. L. Muskens built a model with 3D CNNs to predict the near-field and far-field electric/magnetic response of arbitrary nanostructures [81]. They pixelated the dielectric or plasmonic nanostructure of interest into a 3D image and fed the image into several layers of 3D CNNs. Then an output 3D image with the same size as the input was predicted, representing the electric field under a fixed wavelength and polarization in the same coordination system as shown in Fig. 7(b). The residual connections and shortcut connections in the network are known as the residual learning [82] and U-Net [83] blocks, which can help to stabilize the gradient of the networks and make the network deeper without compromising its performance [84,85]. From the predicted near-field response, other physical quantities, such as far-field scattering patterns, energy flux, and electromagnetic chirality, can then be deduced. The authors studied two cases: 2D gold nanostructures with random polygonal shapes and 3D silicon structures consisting of several pillars. Each scheme was trained by simulation data of 30,000 distinct geometries. With the well-trained model, the authors reproduced several nano-optical effects from the near-field prediction from the 3D CNNs, like antenna behavior of gold nanorods and Kerker-type scattering of Si nanoblocks. The model can potentially serve as an extremely fast tool to replace the current full-wave simulation methods, with the trade-off of slightly decreased accuracy.

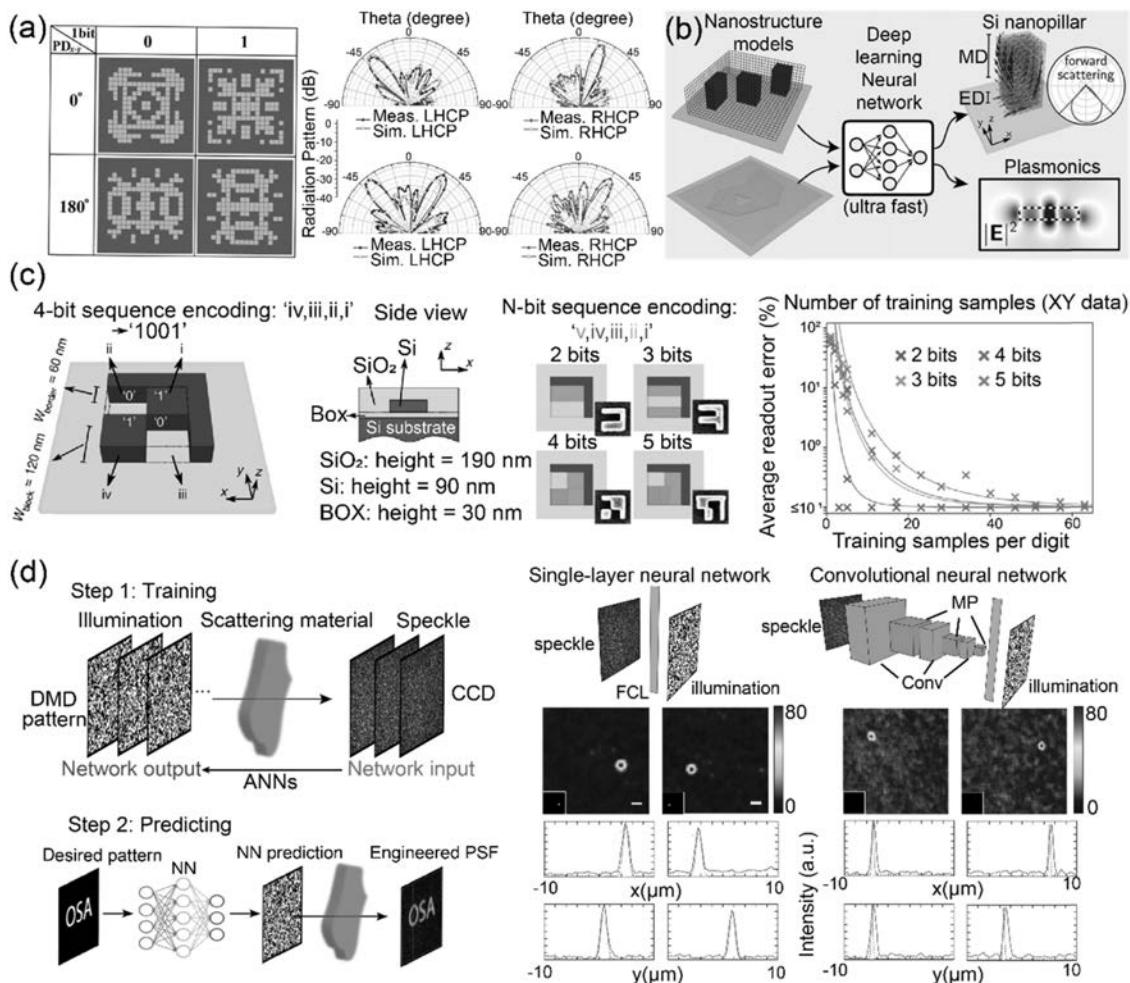
In parallel, a one-dimensional (1D) CNN was also introduced to analyze the scattering spectra of silicon nanostructures for optical information storage as demonstrated by P. R. Wiecha *et al.* in 2019 [86]. The authors used Si nanostructures to store the bit information with high density as shown in the left panel of Fig. 7(c). The nanostructure was divided into  $N$  parts. If a certain part contained a silicon block, the particular

bit was defined as “1;” otherwise it was “0.” Therefore, an  $N$ -bit information storage unit was created. The readout of the information encoded in the nanostructure was through far-field measurement. Here, the dark-field spectra under  $x$ - and  $y$ -polarized light in the visible range were chosen to be the measured information. The 1D CNNs together with FCLs were used to analyze the spectra, where the input of the classification problem was the scattering spectra and the output was the index of the class number among the total  $2^N$  classes for  $N$  bits, representing the bit sequence. The network was trained with experimentally measured dark-field spectra of 625 fabricated nanostructures for each geometry. The model trained after 100 epochs can show quasi-error-free prediction with accuracy higher than 99.97% for the 2-bit to 5-bit (or even 9-bit) geometries as demonstrated in the right panel of Fig. 7(c). The authors further showed that the input information can be greatly reduced by feeding the network with only a small spectral window around 100 nm or even several discrete data points on the spectra, while the effect on the accuracy was neglectable. Finally, the authors managed to retrieve the stored information from the RGB value of the dark-field color image of the nanostructures. This new approach can reduce the complexity and equipment cost of the readout process and at the same time promises a massively parallel retrieval of information.

CNNs are not always the best choice for image inputs as found by A. Turpin *et al.* in 2018 [87]. The scheme of this work is shown on the left of Fig. 7(d). They studied the speckle of the illuminated digital micromirror device (DMD) pattern after light passed through a layer of scattering material like a glass diffuser of multimode fibers. They intended to inversely design the required DMD pattern for an output speckle to form a certain image. The authors built two models by a single FCL and multilayer CNNs. The right panel of Fig. 7(d) presents the result of the inverse designs for the desired Gaussian beam outputs based on the two models. We can find that the measured results of the single FCL look better than those of the multilayer CNNs. Quantitatively, both of the models can achieve a signal-to-noise ratio larger than 10. However, the enhancement metric is  $\eta = 32$  for the first model and only 3.6 for the second model, where  $\eta$  is defined as the intensity at the generated focal point divided by the mean intensity of the background speckle. Therefore, the authors concluded that in this particular application, CNNs can reduce the number of network parameters by almost 80% compared to the single FCL, but at the cost of a worse performance when the used training data have a similar number. The well-trained model can then be used to predict the required illumination pattern with varied output images. In this way, the authors achieved a dynamic scan of the focal point by manipulating the input illumination with a high frame rate of 22.7 kHz.

#### 4. OTHER INTELLIGENT ALGORITHMS FOR PHOTONIC DESIGNS

There are other well-developed computational methods and algorithms that can be applied for the inverse design with satisfactory performance in specific circumstances. One of the most popular methods is the genetic algorithm [88,89], which is inspired by Charles Darwin's natural evolution theory. As



**Fig. 7.** (a) Left: One example of 1-bit coding elements with regular phase differences. Right: Comparison of the simulated and measured results of the dual- and triple-beam coding metasurfaces. (b) Schematic of the proposed 3D CNN model to characterize the near-field and far-field properties of arbitrary dielectric and plasmonic nanostructures. (c) Left: Sketch of the nanostructure geometry and the 1D CNN-based ANNs. Right: Training convergence and readout accuracy of the ANNs. (d) Left: The workflow of designing the DMD pattern for light control through scattering media with ANNs. Right: The structures of the FCLs-based single-layer neural network and the CNNs, together with the simulated and measured results for the focusing effect. (a) is reproduced from Ref. [80] with permission; (b) is reproduced from Ref. [81] with permission; (c) is reproduced from Ref. [86] with permission; (d) is reproduced from Ref. [87] with permission.

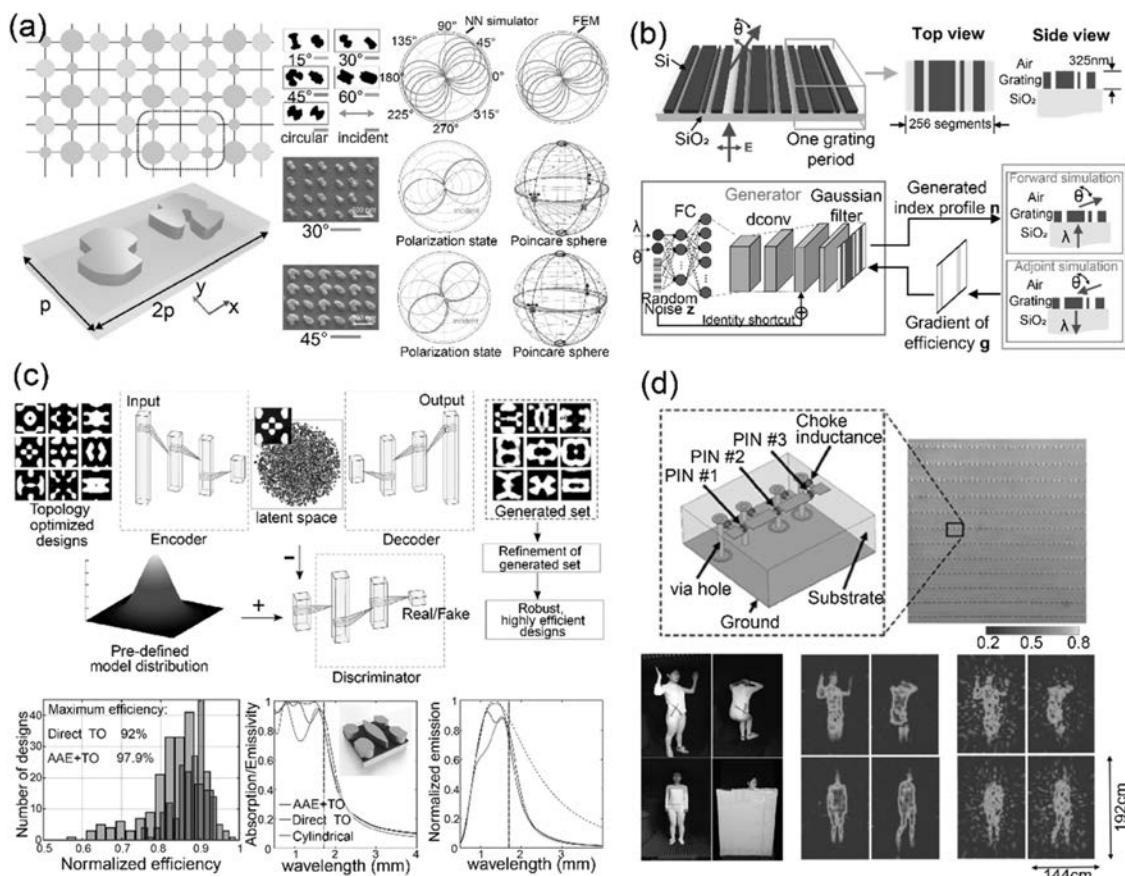
previously discussed [79], in the design toward a target response, a group of initial designs is created either randomly or empirically. The performance of the first generation of “species” is tested and compared to the target response, and a fitness score based on the comparison is calculated. The algorithm will select several “species” in the current generation that has the highest fitness score. Then reproduction combining the information of two or more designs and mutation that adds random noise to the design is performed to generate the next generation of species. The process is repeated until all or most of the species in the new generation have good fitness scores. This algorithm was already applied to photonic design problems a decade ago and achieved great success [90–94]. Recently, Z. Liu *et al.* published their work that integrated the genetic algorithm with ANNs [95]. They studied “meta-molecules” consisting of multiple meta-atoms that can realize polarization conversion and anomalous light deflection as shown on the left of Fig. 8(a).

The model is composed of a compositional pattern-producing network (CPPN), which is used to decode the 2D patterns from a latent variable, and a cooperative coevolution algorithm (CC) to identify a set of vectors in the latent space. The CPPNs take the coordinate tuple  $(x_i, y_i, r_i)$  one at a time together with a latent vector  $v$ , which controls the shapes of the patterns, and assemble the predictions from the whole input as a pattern. The CC then performs the genetic algorithm with the fitness score calculated based on the output polarization state, the ellipticity, and the phase and intensity of the electric field. The authors first trained a neural network simulator with the response from 8000 meta-atoms in different shapes. This simulator can be adopted in the CC to greatly reduce the time of fitness score computation. The simulator can achieve predictions of real and imaginary parts of spectra with an accuracy above 97%. The authors designed and fabricated meta-molecules comprising two (or eight) meta-atoms to implement polarization conversion

under linear polarization as well as anomalous light deflection under circular polarization. The simulated and measured results of polarization conversion are plotted in the right panel of Fig. 8(a), showing excellent agreement with the target.

Another widely used optimization algorithm for the inverse design is gradient-based topology optimization [21,96–103]. In the optimization process, the design space is discretized into pixels whose properties (i.e., refractive index) can be represented by a parameter set  $p$ . The parameter set will be optimized for a prescribed target response by maximizing (minimizing) a user-defined objective function  $\mathcal{F}$ . Starting from an initial parameter set, both a forward simulation and an adjoint simulation are performed to calculate the gradient of the objective function  $\partial\mathcal{F}/\partial p_i$  with respect to each parameter. Then the parameters are updated according to the gradient ascent (descent) method. This iterative process is continued until the objective function is well optimized. Taking advantage of the topology optimization, J. Jiang *et al.* presented a global optimizer for highly efficient metasurfaces that can deflect light to desired angles [100]. As illustrated in the top panel of Fig. 8(b), the metagrating in one period is divided into 256 segments, and

each segment can be filled with either air or Si. To optimize the metagrating, the authors used a global optimization method named GLOnet. The GLOnet is based on both a generative neural network (GNN) and topology optimization as shown in the bottom panel of Fig. 8(b). The GNN takes the desired deflection angle  $\theta$  and the working wavelength  $\lambda$  together with a random noise vector  $z$  as inputs. The inputs pass through FCLs and layers of deconvolutional blocks, and then a metagrating design is generated. The Gaussian filter at the last layer of the generator eliminates small features that are hard to fabricate. Next, the topology optimization is applied. By performing both a forward simulation and an adjoint simulation, the gradient of the objective function (efficiency) is calculated. The weights of the ANNs are updated according to the gradient ascent method. To make the model capable of working for any deflection angle and wavelength, the initialization of the model is essential to span the full design space. Therefore, an identity shortcut is added to map the random noise directly to the output design, which will enable all kinds of designs when the initial weight of the GNN is small. It should be noted that the GLOnet is different from conventional topology



**Fig. 8.** (a) Left: Illustration of meta-molecules. Right: Fabricated samples and the measured and simulated results of polarization conversion. (b) Top: Schematic of a silicon metagrating that deflects light to a certain angle. Bottom: The proposed conditional GLOnet for metagrating optimization. (c) Top: Schematic of structure refinement and filtering for the high-efficiency thermal emitter. Bottom: The efficiency, emissivity, and normalized emission of the well-optimized thermal emitter. (d) Top: Illustration of the unit cell consisting of three metallic patches connected via PIN diodes and a photograph of the fabricated metasurface. Bottom: Experimental results for reconstructing human body imaging. (a) is reproduced from Ref. [95] with permission; (b) is reproduced from Ref. [100] with permission; (c) is reproduced from Ref. [42] with permission; (d) is reproduced from Ref. [104] with permission.

optimization. In conventional topology optimization, the structural parameters (like the refractive index of individual segments) are updated for a single device with a fixed deflection angle and wavelength. When the goal (deflection angle  $\theta$ ) or the working wavelength is changed, the optimization needs to be performed again for the new device. However, in the GLOnet, the optimized parameters are the weights in the neural networks during each iteration. Therefore, the GNN is improved in terms of the ability to inversely design devices for varied goals and working wavelengths, without the need to retrain the model when the target changes. The performances of conventional topology optimization and the GLOnet optimization have been compared in this work: 92% of the devices designed by the GLOnet have efficiencies higher than or within 5% of the devices designed by the other method. In addition, the retrieved devices gradually converge to a high-efficiency region as the iteration number of the training process increases.

Combining topology optimization and ANNs, Z. A. Kudyshev *et al.* studied the structure optimization of high-efficiency thermophotovoltaic (TPV) cells operating in the desired wavelength range ( $\lambda = 0.5\text{--}1.7 \mu\text{m}$ ) [42]. The design is based on a gap plasmonic structure. As shown in the top panel of Fig. 8(c), the optimization can be divided into three main steps. First, the topology optimization method is applied to generate a group of appropriate structures for training. Then an adversarial autoencoder (AAE) network is trained. Similar to the VAE, the AAE consists of an encoder to map the input designs to a latent space and a decoder to retrieve the structure from the latent vector sampled from the latent space. Both the VAE and AAE models try to make the latent distribution  $q(\tilde{z})$  approach a predefined distribution  $p(z)$  (a 15-dimensional Gaussian distribution in Ref. [42]). In the VAE model, a Kullback–Leibler divergence that compares  $q(\tilde{z})$  with  $p(z)$  is defined as one part of the loss function; while in the AAE, a discriminator used to distinguish the samples from  $q(\tilde{z})$  and  $p(z)$  is built, and the encoder is trained to generate samples that can fool the discriminator. In the last step, the structure retrieved from the decoder is refined with topology optimization to remove the blurring of the generated designs. As a result, the hybrid method that combines AAE and topology optimization shows great performance, providing a mean efficiency of 90% for the retrieved structures. In contrast, the efficiency is 82% via direct topology optimization. The comparison between these two methods is shown at the bottom of Fig. 8(c) together with the emissivity and emission plots for the best designs from either method. In a very recent work [105], the same group further developed a global optimization method in which a global optimization engine can generate latent vectors and Visual Geometry Groupnet can rapidly assess the performance of the design.

Conventional machine learning methods, such as Bayesian learning [106], clustering [107], and manifold learning [104], are also very helpful in solving photonic design problems. In 2019, L. Li *et al.* showcased a machine-learning-based imager that can efficiently record the microwave image of a moving object by a reprogrammable metasurface [104]. This work may pave the way for intelligent surveillance with both fast response time and high accuracy. The meta-atom has three

metallic patches connected via PIN diodes to encode 2-bit information as schematically shown in the top panel of Fig. 8(d). The digital phase step is around  $90^\circ$  between adjacent states, and the state can be tuned by applying an external bias voltage. The authors recorded a moving person for less than 20 min to generate the training data for the model. With principal component analysis (or random projection), the main modes with significant contributions were calculated. Then all meta-atoms were tuned by a bias voltage to match the principal component analysis modes for each measurement. In this way, the measurement became more efficient because it always captured the information with a high contribution to reconstructing the microwave image. To test the well-trained model, another person was moving in front of the metasurface, and images of the movements were reconstructed as shown at the bottom of Fig. 8(d). With only 400 measurements, which were far fewer than the number of pixels, high-quality images could be produced even when the person was blocked by a 3-cm-thick paper wall. This method was further extended to the classification problem, in which the authors defined three different movements (i.e., standing, bending, and raising arms). With a simple nearest-neighbor algorithm, only 25 measurements led to good recognition of the movements.

## 5. CONCLUSION AND OUTLOOK

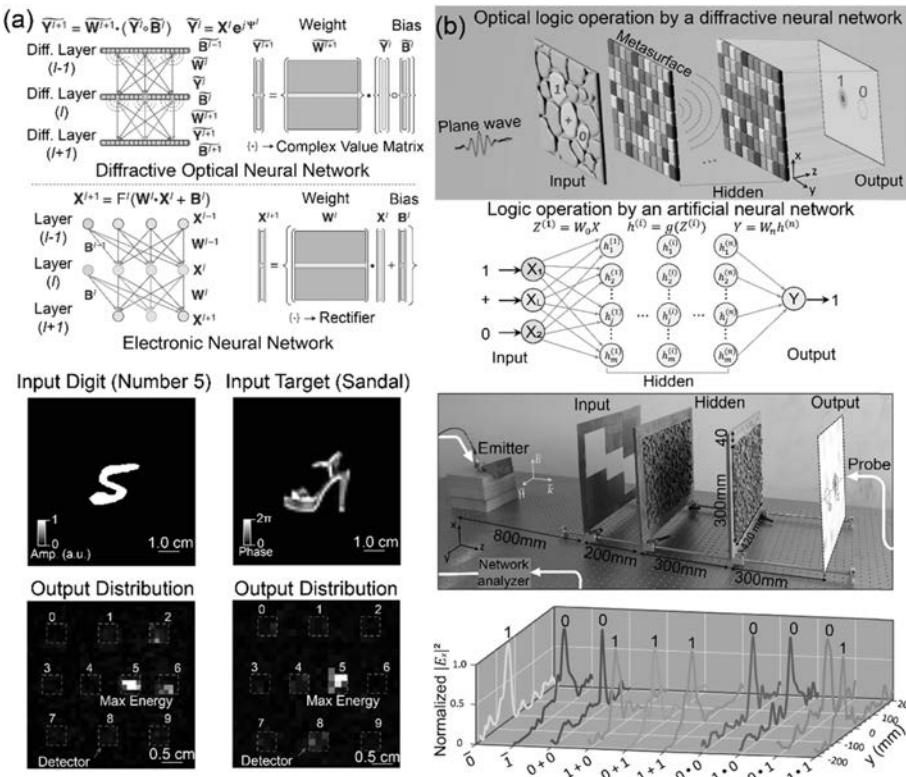
In this review, we have introduced the basic idea of applying ANNs and other advanced algorithms to accelerate and optimize photonic designs, including plasmonic nanostructures and metamaterials. We have highlighted some representative works in this field and discussed the performance and applications of the proposed models. In the inverse design problem, the neural network is usually built upon FCLs and CNNs, integrated with other neural network units like ResNets and RNNs. It is beneficial to incorporate ANNs with conventional optimization methods such as genetic algorithm and topology optimization because the conventional optimization methods can help to perform global optimization and provide feedback to further improve the ANNs. The emergence of all the methods offers a great opportunity to increase the structural complexity in the devices, which can realize much more complex and novel functionalities.

The development of photonics can also potentially benefit the studies of computational methods. For instance, it has been long sought to push the computation speed to the speed of the light. All-optical neuromorphic computing [108–112] via optical networks is one approach toward this goal. In principle, the diffraction nature of light described as  $\exp(i\vec{k} \cdot \vec{r})$  can also be regarded as a nonlinear function. Therefore, the intensity profiles in two diffractive layers “connected” with light diffraction can be a good analogy to the connection between neurons in ANNs. Based on this idea, researchers have demonstrated a new kind of neural network built upon all-optical components, which are known as optical neural networks (ONNs) [113–117]. As a comprehensive example, X. Lin *et al.* reported an all-optical system that can serve as a diffractive deep neural network (D<sup>2</sup>NN) for image classification in 2018 [118]. The system is composed of several layers of 3D printed structures. According to the Huygens–Fresnel principle, points in the

$D^2NN$  layers can be regarded as a secondary source of light. Therefore, each point in the front layer will contribute to the amplitude and phase distribution of each point in the following layers, while the propagation phase will function as the nonlinearity. The analogy between the  $D^2NN$  and the ANN is illustrated in the top panel of Fig. 9(a). The authors designed the  $D^2NN$  using the same error backpropagation method as in the ANNs and adjusted the phase distribution in each layer. This design process was run on the computer, but once the design finished, the fabricated device can perform prediction (classification) all-optically. In the measurement, the light passed through an input plane with the same shape as the image. By detecting the position with the maximum output intensity after light passing through all layers, the class of the input image can be read out. The authors trained and tested the classifier with images of handwritten digits and fashion products. The experimental results show great accordance with the expectation, as shown in the bottom panel of Fig. 9(a), with an accuracy of 91.75% and 86.60% for the two tasks, respectively. Two years later, C. Qian *et al.* showed optical logic operations by a diffractive neural network [119]. The goal was to perform the logic operations such as “and,” “or,” and “not” for the inputs. As shown in the first two rows of Fig. 9(b), the input wave was shaped so that it can only pass through certain regions before illuminating on the diffractive metasurface. In this way, the two binary inputs and the logical operation can be controlled. The results can also be read out by detecting the intensity at two positions representing “0” and “1.” The last two

rows of Fig. 9(b) show the experimental measurement for 10 different operations, and all the profiles indicate the correct results. More efforts are needed to further advance this exciting direction, for instance, by reducing the footprint and increasing the efficiency of the optical neural networks.

The ANNs are typically considered a “black box” since the relationship between inputs and outputs learned by the ANNs is usually implicit. In some published works, researchers can visualize the output of each individual layer to provide some information on what feature is learned (or what function is done) by each layer [40], which is a good attempt. However, if we can further extract the relation explicitly from the well-trained ANNs, it will be very helpful to find new structure groups that lie out of the conventional geometry groups (like H-shape, C-shape, bowtie). At the same time, it will also provide guidelines or insights for the design of optical devices. Another important direction is to extend the generality of the ANNs models. When applying ANNs to solve the traditional tasks, such as image recognition and natural language processing, we want the neural networks to learn the information and distribution that lie inside the natural images or languages themselves and try to reconstruct or approximate these distributions. The ANNs have been proven to work well in learning and summarizing the distributions from the images or languages. At the same time, it is relatively easy to extend the model to deal with other kinds of images or languages. However, the inverse design tasks in photonics are more complicated. The reason is that the ANNs need to learn the



**Fig. 9.** (a) Top: Comparison between the all-optical  $D^2NN$  and a conventional ANN. Bottom: Measured performance of the classifier for handwritten digits and fashion products. (b) Top: Sketch of the optical logic operations by a diffractive neural network. Bottom: Experiment setup and measured results of three basic logic operations on the fabricated metasurface. (a) is reproduced from Ref. [118] with permission; (b) is reproduced from Ref. [119] with permission.

implicit physical rules (such as Maxwell's equations) between the structures and their optical responses, instead of the information and distribution associated with the structures themselves. Therefore, extending the capability of a well-trained neural network in the inverse design problems remains a challenge. Most of the ANNs described in this review paper are only specified for a certain design platform or application. It is true that a model can be fine-tuned to handle different tasks, but the model needs to be retrained and, at the same time, an additional training data set is required. When the original training set contains all kinds of training data for multiple tasks, multiple design rules are likely to be involved and learned by the ANNs. The performance of the model will not be satisfactory for each individual task compared to the model trained with only a specific data set for this task, because the rules for other tasks will serve as perturbation or noise in this case. It is very important to find the trade-off.

Over the past decades, photonics and artificial intelligence have been evolving largely as two separate research disciplines. The intersection and combination of these two topics in recent years have brought exciting achievements. On one hand, the innovative ANN models provide a powerful tool to accelerate the optical design and implementation process. Some nonintuitive structures and phenomena have been discovered by this new strategy. On the other hand, the developed optical designs are expected to produce a variety of real-world applications, such as optical imaging, holography, communications, and information encryption, with high efficiency, fidelity, and robustness. Toward this goal, we need to include the practical fabrication constraints and underlying material properties into the design space in order to globally optimize the devices and systems. We believe that the field of interfacing photonics and artificial intelligence will significantly move forward as more researchers from different backgrounds join this effort.

**Funding.** National Science Fundation (ECCS-1916839).

**Disclosures.** The authors declare no conflicts of interest to this paper.

## REFERENCES

- Y. Liu and X. Zhang, "Metamaterials: a new frontier of science and technology," *Chem. Soc. Rev.* **40**, 2494–2507 (2011).
- N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *Science* **334**, 333–337 (2011).
- Y. W. Huang, H. W. H. Lee, R. Sokhoyan, R. A. Pala, K. Thyagarajan, S. Han, D. P. Tsai, and H. A. Atwater, "Gate-tunable conducting oxide metasurfaces," *Nano Lett.* **16**, 5319–5325 (2016).
- A. M. Shaltout, V. M. Shalaev, and M. L. Brongersma, "Spatiotemporal light control with active metasurfaces," *Science* **364**, eaat3100 (2019).
- L. Li, K. Yao, Z. Wang, and Y. Liu, "Harnessing evanescent waves by bianisotropic metasurfaces," *Laser Photon. Rev.* **14**, 1900244 (2020).
- L. Huang, X. Chen, H. Mühlenernd, H. Zhang, S. Chen, B. Bai, Q. Tan, G. Jin, K.-W. Cheah, C.-W. Qiu, J. Li, T. Zentgraf, and S. Zhang, "Three-dimensional optical holography using a plasmonic metasurface," *Nat. Commun.* **4**, 2808 (2013).
- G. Zheng, H. Mühlenernd, M. Kenney, G. Li, T. Zentgraf, and S. Zhang, "Metasurface holograms reaching 80% efficiency," *Nat. Nanotechnol.* **10**, 308–312 (2015).
- B. Wang, F. Dong, Q. T. Li, D. Yang, C. Sun, J. Chen, Z. Song, L. Xu, W. Chu, Y. F. Xiao, and Q. Gong, "Visible-frequency dielectric metasurfaces for multiwavelength achromatic and highly dispersive holograms," *Nano Lett.* **16**, 5235–5240 (2016).
- L. Jin, Z. Dong, S. Mei, Y. F. Yu, Z. Wei, Z. Pan, S. D. Rezaei, X. Li, A. I. Kuznetsov, Y. S. Kivshar, and J. K. Yang, "Noninterleaved metasurface for (2<sup>6</sup>–1) spin-and wavelength-encoded holograms," *Nano Lett.* **18**, 8016–8024 (2018).
- X. Chen, L. Huang, H. Mühlenernd, G. Li, B. Bai, Q. Tan, G. Jin, C. W. Qiu, S. Zhang, and T. Zentgraf, "Dual-polarity plasmonic metalenses for visible light," *Nat. Commun.* **3**, 1198 (2012).
- X. Ni, S. Ishii, A. V. Kildishev, and V. M. Shalaev, "Ultra-thin, planar, Babinet-inverted plasmonic metalenses," *Light Sci. Appl.* **2**, e72 (2013).
- S. Wang, P. C. Wu, V. C. Su, Y. C. Lai, M. K. Chen, H. Y. Kuo, B. H. Chen, Y. H. Chen, T. T. Huang, J. H. Wang, and R. M. Lin, "A broadband achromatic metalens in the visible," *Nat. Nanotechnol.* **13**, 227–232 (2018).
- W. T. Chen, A. Y. Zhu, V. Sanjeev, M. Khorasaninejad, Z. Shi, E. Lee, and F. Capasso, "A broadband achromatic metalens for focusing and imaging in the visible," *Nat. Nanotechnol.* **13**, 220–226 (2018).
- X. Zang, H. Ding, Y. Intaravanne, L. Chen, Y. Peng, J. Xie, Q. Ke, A. V. Balakin, A. P. Shkurinov, X. Chen, and Y. Zhu, "A multi-foci metalens with polarization-rotated focal points," *Laser Photon. Rev.* **13**, 1900182 (2019).
- M. Faraji-Dana, E. Arbabi, A. Arbabi, S. M. Kamali, H. Kwon, and A. Faraon, "Compact folded metasurface spectrometer," *Nat. Commun.* **9**, 4196 (2018).
- A. Tittl, A. Leitis, M. Liu, F. Yesilkoy, D. Y. Choi, D. N. Neshev, Y. S. Kivshar, and H. Altug, "Imaging-based molecular barcoding with pixelated dielectric metasurfaces," *Science* **360**, 1105–1109 (2018).
- A. Leitis, A. Tittl, M. Liu, B. H. Lee, M. B. Gu, Y. S. Kivshar, and H. Altug, "Angle-multiplexed all-dielectric metasurfaces for broadband molecular fingerprint retrieval," *Sci. Adv.* **5**, eaaw2871 (2019).
- N. K. Grady, J. E. Heyes, D. R. Chowdhury, Y. Zeng, M. T. Reiten, A. K. Azad, A. J. Taylor, D. A. Dalvit, and H. T. Chen, "Terahertz metamaterials for linear polarization conversion and anomalous refraction," *Science* **340**, 1304–1307 (2013).
- M. Kim, K. Yao, G. Yoon, I. Kim, Y. Liu, and J. Rho, "A broadband optical diode for linearly polarized light using symmetry-breaking metamaterials," *Adv. Opt. Mater.* **5**, 1700600 (2017).
- L. Kang, S. P. Rodrigues, M. Taghinejad, S. Lan, K. T. Lee, Y. Liu, D. H. Werner, A. Urbas, and W. Cai, "Preserving spin states upon reflection: linear and nonlinear responses of a chiral meta-mirror," *Nano Lett.* **17**, 7102–7109 (2017).
- Z. Shi, A. Y. Zhu, Z. Li, Y. W. Huang, W. T. Chen, C. W. Qiu, and F. Capasso, "Continuous angle-tunable birefringence with freeform metasurfaces for arbitrary polarization conversion," *Sci. Adv.* **6**, eaaba3367 (2020).
- Z. Wang, K. Yao, M. Chen, H. Chen, and Y. Liu, "Manipulating Smith-Purcell emission with babinet metasurfaces," *Phys. Rev. Lett.* **117**, 157401 (2016).
- D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies," *Science* **314**, 977–980 (2006).
- J. K. Gansel, M. Thiel, M. S. Rill, M. Decker, K. Bade, V. Saile, G. von Freymann, S. Linden, and M. Wegener, "Gold helix photonic metamaterial as broadband circular polarizer," *Science* **325**, 1513–1515 (2009).
- X. Liu, T. Tyler, T. Starr, A. F. Starr, N. M. Jokerst, and W. J. Padilla, "Taming the blackbody with infrared metamaterials as selective thermal emitters," *Phys. Rev. Lett.* **107**, 045901 (2011).
- D. P. Fromm, A. Sundaramurthy, P. J. Schuck, G. Kino, and W. Moerner, "Gap-dependent optical coupling of single "bowtie" nanoantennas resonant in the visible," *Nano Lett.* **4**, 957–961 (2004).

27. M. Choi, S. H. Lee, Y. Kim, S. B. Kang, J. Shin, M. H. Kwak, K. Y. Kang, Y. H. Lee, N. Park, and B. Min, "A terahertz metamaterial with unnaturally high refractive index," *Nature* **470**, 369–373 (2011).
28. S. Sun, Q. He, S. Xiao, Q. Xu, X. Li, and L. Zhou, "Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves," *Nat. Mater.* **11**, 426–431 (2012).
29. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**, 339–366 (2019).
30. Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, "Artificial neural networks enabled by nanophotonics," *Light Sci. Appl.* **8**, 42 (2019).
31. R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nano. Adv.* **2**, 1007–1023 (2020).
32. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **9**, 1041–1057 (2020).
33. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2020).
34. M. M. R. Elsayy, S. Lanteri, R. Duvigneau, J. A. Fan, and P. Genevet, "Numerical optimization methods for metasurfaces," *Laser Photon. Rev.* **14**, 1900445 (2020).
35. D. Piccinotti, K. F. MacDonald, S. Gregory, I. Youngs, and N. I. Zheludev, "Artificial intelligence for photonics and photonic materials," *Rep. Prog. Phys.* **84**, 012401 (2020).
36. L. Huang, L. Xu, and A. E. Miroshnichenko, "Deep learning enabled nanophotonics," in *Advances and Applications in Deep Learning* (IntechOpen, 2020).
37. J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *Nat. Rev. Mater.* (2020).
38. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
39. T. Qiu, X. Shi, J. Wang, Y. Li, S. Qu, Q. Cheng, T. Cui, and S. Sui, "Deep learning: a rapid and efficient route to automatic metasurface design," *Adv. Sci.* **6**, 1900128 (2019).
40. I. Sajedian, J. Kim, and J. Rho, "Finding the optical properties of plasmonic structures by image processing using a combination of convolutional neural networks and recurrent neural networks," *Microsyst. Nanoeng.* **5**, 27 (2019).
41. W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Adv. Mater.* **31**, 1901111 (2019).
42. Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine-learning-assisted metasurface design for high-efficiency thermal emitter optimization," *Appl. Phys. Rev.* **7**, 021407 (2020).
43. F. Wen, J. Jiang, and J. A. Fan, "Robust freeform metasurface design based on progressively growing generative networks," *ACS Photon.* **7**, 2098–2104 (2020).
44. W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural Netw.* **10**, 1659–1671 (1997).
45. W. Gerstner and W. M. Kistler, *Spiking Neuron Models: Single Neurons, Populations, Plasticity* (Cambridge University, 2002).
46. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
47. S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core-shell nanoparticles," *ACS Appl. Mater. Interfaces* **11**, 24264–24268 (2019).
48. Y. Chen, J. Zhu, Y. Xie, N. Feng, and Q. H. Liu, "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," *Nanoscale* **11**, 9749–9755 (2019).
49. Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljačić, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photon.* **6**, 1168–1174 (2019).
50. I. Sajedian, T. Badloe, H. Lee, and J. Rho, "Deep Q-network to produce polarization-independent perfect solar absorbers: a statistical report," *Nano Converg.* **7**, 26 (2020).
51. I. Sajedian, H. Lee, and J. Rho, "Double-deep Q-learning to increase the efficiency of metasurface holograms," *Sci. Rep.* **9**, 10899 (2019).
52. I. Sajedian, H. Lee, and J. Rho, "Design of high transmission color filters for solar cells directed by deep Q-learning," *Sol. Energy* **195**, 670–676 (2020).
53. H. Wang, Z. Zheng, C. Ji, and L. J. Guo, "Automated multi-layer optical design via deep reinforcement learning," *Mach. Learn.* **2**, 025013 (2020).
54. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
55. T. Zhang, J. Wang, Q. Liu, J. Zhou, J. Dai, X. Han, Y. Zhou, and K. Xu, "Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks," *Photon. Res.* **7**, 368–380 (2019).
56. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Opt. Express* **27**, 27523–27535 (2019).
57. G. Alagappan and C. E. Png, "Deep learning models for effective refractive indices in silicon nitride waveguides," *J. Opt.* **21**, 035801 (2019).
58. M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.* **9**, 1368 (2019).
59. Y. Long, J. Ren, Y. Li, and H. Chen, "Inverse design of photonic topological state via machine learning," *Appl. Phys. Lett.* **114**, 181105 (2019).
60. L. Pilozzi, F. A. Farrelly, G. Marcucci, and C. Conti, "Machine learning inverse problem for topological photonics," *Commun. Phys.* **1**, 57 (2018).
61. G. Alagappan and C. E. Png, "Modal classification in optical waveguides using deep learning," *J. Mod. Opt.* **66**, 557–561 (2018).
62. I. Sajedian, T. Badloe, and J. Rho, "Optimization of colour generation from dielectric nanostructures using reinforcement learning," *Opt. Express* **27**, 5874–5883 (2019).
63. H. Ren, W. Shao, Y. Li, F. Salim, and M. Gu, "Three-dimensional vectorial holography based on machine learning inverse design," *Sci. Adv.* **6**, eaaz4261 (2020).
64. C. Qian, B. Zheng, Y. Shen, L. Jing, E. Li, L. Shen, and H. Chen, "Deep-learning-enabled self-adaptive microwave cloak without human intervention," *Nat. Photonics* **14**, 383–390 (2020).
65. Z. Liu, D. Zhu, S. P. Rodrigues, K. T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
66. J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Free-form diffractive metagrating design based on generative adversarial networks," *ACS Nano* **13**, 8872–8878 (2019).
67. R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Sci. Rep.* **9**, 19728 (2019).
68. T. Zahavy, A. Dikopoltsev, D. Moss, G. I. Haham, O. Cohen, S. Mannor, and M. Segev, "Deep learning reconstruction of ultrashort pulses," *Optica* **5**, 666–673 (2018).
69. S. So and J. Rho, "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**, 1255–1261 (2019).
70. J. A. Schuller, E. S. Barnard, W. Cai, Y. C. Jun, J. S. White, and M. L. Brongersma, "Plasmonics for extreme light concentration and manipulation," *Nat. Mater.* **9**, 193–204 (2010).
71. D. K. Gramotnev and S. I. Bozhevolnyi, "Plasmonics beyond the diffraction limit," *Nat. Photonics* **4**, 83–91 (2010).
72. J. Lin, J. B. Mueller, Q. Wang, G. Yuan, N. Antoniou, X. C. Yuan, and F. Capasso, "Polarization-controlled tunable directional coupling of surface plasmon polaritons," *Science* **340**, 331–334 (2013).
73. Z. Cai, Y. Xu, C. Wang, and Y. Liu, "Polariton photonics using structured metals and 2D materials," *Adv. Opt. Mater.* **8**, 1901090 (2019).
74. Y. Liu, S. Palomba, Y. Park, T. Zentgraf, X. Yin, and X. Zhang, "Compact magnetic antennas for directional excitation of surface plasmons," *Nano Lett.* **12**, 4853–4858 (2012).
75. I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial

- networks," in *Advances in Neural Information Processing Systems* (2014), pp. 2672–2680.
76. D. P. Kingma and M. Welling, "Auto-encoding variational Bayes," in *2nd International Conference on Learning Representations (ICLR)* (2014), pp. 1–14.
  77. S. J. Wetzel, "Unsupervised learning of phase transitions: from principal component analysis to variational autoencoders," *Phys. Rev. E* **96**, 022140 (2017).
  78. W. Ma and Y. Liu, "A data-efficient self-supervised deep learning model for design and characterization of nanophotonic structures," *Sci. China Phys. Mech. Astron.* **63**, 284212 (2020).
  79. Z. Liu, L. Raju, D. Zhu, and W. Cai, "A hybrid strategy for the discovery and design of photonic structures," *IEEE J. Emerg. Sel. Top. Circuits Syst.* **10**, 126–135 (2020).
  80. Q. Zhang, C. Liu, X. Wan, L. Zhang, S. Liu, Y. Yang, and T. J. Cui, "Machine-learning designs of anisotropic digital coding metasurfaces," *Adv. Theor. Simul.* **2**, 1800132 (2018).
  81. P. R. Wiecha and O. L. Muskens, "Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures," *Nano Lett.* **20**, 329–338 (2020).
  82. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
  83. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), pp. 234–241.
  84. D. Balduzzi, M. Frean, L. Leary, J. P. Lewis, K. W. Ma, and B. McWilliams, "The shattered gradients problem: if resnets are the answer, then what is the question?" in *Proceedings of the 34th International Conference on Machine Learning* (2017), pp. 342–350.
  85. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: inverted residuals and linear bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 4510–4520.
  86. P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nat. Nanotechnol.* **14**, 237–244 (2019).
  87. A. Turpin, I. Vishniakou, and J. D. Seelig, "Light scattering control in transmission and reflection with neural networks," *Opt. Express* **26**, 30911–30929 (2018).
  88. D. E. Goldberg and J. H. Holland, *Genetic Algorithms and Machine Learning* (Springer, 1988).
  89. J. H. Holland, *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence* (MIT, 1992).
  90. L. Shen, Z. Ye, and S. He, "Design of two-dimensional photonic crystals with large absolute band gaps using a genetic algorithm," *Phys. Rev. B* **68**, 035109 (2003).
  91. S. Preble, M. Lipson, and H. Lipson, "Two-dimensional photonic crystals designed by evolutionary algorithms," *Appl. Phys. Lett.* **86**, 061111 (2005).
  92. T. Feichtner, O. Selig, M. Kiunke, and B. Hecht, "Evolutionary optimization of optical antennas," *Phys. Rev. Lett.* **109**, 127701 (2012).
  93. C. Wang, S. Yu, W. Chen, and C. Sun, "Highly efficient light-trapping structure design inspired by natural evolution," *Sci. Rep.* **3**, 1025 (2013).
  94. M. D. Huntington, L. J. Lauhon, and T. W. Odom, "Subwavelength lattice optics by evolutionary design," *Nano Lett.* **14**, 7195–7200 (2014).
  95. Z. Liu, D. Zhu, K. T. Lee, A. S. Kim, L. Raju, and W. Cai, "Compounding meta-atoms into metamolecules with hybrid artificial intelligence techniques," *Adv. Mater.* **32**, 1904790 (2020).
  96. P. I. Borel, A. Harpøth, L. H. Frandsen, M. Kristensen, P. Shi, J. S. Jensen, and O. Sigmund, "Topology optimization and fabrication of photonic crystal structures," *Opt. Express* **12**, 1996–2001 (2004).
  97. J. S. Jensen and O. Sigmund, "Topology optimization for nanophotonics," *Laser Photon. Rev.* **5**, 308–321 (2011).
  98. Z. Lin, B. Groever, F. Capasso, A. W. Rodriguez, and M. Lončar, "Topology-optimized multilayered metaoptics," *Phys. Rev. Appl.* **9**, 044030 (2018).
  99. R. Matzen, J. S. Jensen, and O. Sigmund, "Topology optimization for transient response of photonic crystal structures," *J. Opt. Soc. Am. B* **27**, 2040–2050 (2010).
  100. J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.* **19**, 5366–5372 (2019).
  101. T. Phan, D. Sell, E. W. Wang, S. Doshay, K. Edee, J. Yang, and J. A. Fan, "High-efficiency, large-area, topology-optimized metasurfaces," *Light Sci. Appl.* **8**, 48 (2019).
  102. M. Mansoureh, H. Kwon, E. Arbabi, A. McClung, A. Faraon, and A. Arbabi, "Multifunctional 25D metastructures enabled by adjoint optimization," *Optica* **7**, 77–84 (2020).
  103. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
  104. L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C. W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," *Nat. Commun.* **10**, 1082 (2019).
  105. Z. A. Kudryshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine learning-assisted global optimization of photonic devices," *Nanophotonics* **10**, 371–383 (2020).
  106. R. Patel, K. Roy, J. Choi, and K. J. Han, "Generative design of electromagnetic structures through Bayesian learning," *IEEE Trans. Magn.* **54**, 9900138 (2018).
  107. Y. Long, J. Ren, and H. Chen, "Unsupervised manifold clustering of topological phononics," *Phys. Rev. Lett.* **124**, 185501 (2020).
  108. D. Marković, A. Mizrahi, D. Querlioz, and J. Grollier, "Physics for neuromorphic computing," *Nat. Rev. Phys.* **2**, 499–510 (2020).
  109. Y. van de Burgt, A. Melianas, S. T. Keene, G. Malliaras, and A. Salleo, "Organic electronics for neuromorphic computing," *Nat. Electron.* **1**, 386–397 (2018).
  110. J. Torrejon, M. Riou, F. A. Araujo, S. Tsunegi, G. Khalsa, D. Querlioz, P. Bartolotti, V. Cros, K. Yakushiji, A. Fukushima, and H. Kubota, "Neuromorphic computing with nanoscale spintronic oscillators," *Nature* **547**, 428–431 (2017).
  111. D. Brunner, S. Reitzenstein, and I. Fischer, "All-optical neuromorphic computing in optical networks of semiconductor lasers," in *IEEE International Conference on Rebooting Computing (ICRC)* (2016), pp. 1–2.
  112. A. Katumba, M. Freiberger, F. Laporte, A. Lugnan, S. Sackesyn, C. Ma, J. Dambre, and P. Bienstman, "Neuromorphic computing based on silicon photonics and reservoir computing," *IEEE J. Sel. Top. Quantum Electron.* **24**, 8300310 (2018).
  113. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
  114. H. Bagherian, S. Skirlo, Y. Shen, H. Meng, V. Ceperic, and M. Soljacic, "On-chip optical convolutional neural networks," *arXiv:1808.03303* (2018).
  115. J. Bueno, S. Maktoobi, L. Froehly, I. Fischer, M. Jacquot, L. Larger, and D. Brunner, "Reinforcement learning in a large-scale photonic recurrent neural network," *Optica* **5**, 756–760 (2018).
  116. T. Zhang, J. Wang, Y. Dan, Y. Lanqiu, J. Dai, X. Han, X. Sun, and K. Xu, "Efficient training and design of photonic neural network through neuroevolution," *Opt. Express* **27**, 37150–37163 (2019).
  117. E. Khoram, A. Chen, D. Liu, L. Ying, Q. Wang, M. Yuan, and Z. Yu, "Nanophotonic media for artificial neural inference," *Photon. Res.* **7**, 823–827 (2019).
  118. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
  119. C. Qian, X. Lin, X. Lin, J. Xu, Y. Sun, E. Li, B. Zhang, and H. Chen, "Performing optical logic operations by a diffractive neural network," *Light Sci. Appl.* **9**, 59 (2020).

# PHOTONICS Research

## On-demand design of spectrally sensitive multiband absorbers using an artificial neural network

SUNAE SO,<sup>1</sup> YOUNGHWAN YANG,<sup>1</sup> TAEJUN LEE,<sup>1</sup> AND JUNSUK RHO<sup>1,2,3,\*</sup> 

<sup>1</sup>Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea

<sup>2</sup>Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea

<sup>3</sup>National Institute of Nanomaterials Technology (NINT), Pohang 37673, Republic of Korea

\*Corresponding author: jsrho@postech.ac.kr

Received 25 November 2020; revised 12 January 2021; accepted 3 February 2021; posted 5 February 2021 (Doc. ID 415789); published 31 March 2021

We report an approach assisted by deep learning to design spectrally sensitive multiband absorbers that work in the visible range. We propose a five-layered metal-insulator-metal grating structure composed of aluminum and silicon dioxide, and we design its structural parameters by using an artificial neural network (ANN). For a spectrally sensitive design, spectral information of resonant wavelengths is additionally provided as input as well as the reflection spectrum. The ANN facilitates highly robust design of a grating structure that has an average mean squared error (MSE) of 0.023. The optical properties of the designed structures are validated using electromagnetic simulations and experiments. Analysis of design results for gradually changing target wavelengths of input shows that the trained ANN can learn physical knowledge from data. We also propose a method to reduce the size of the ANN by exploiting observations of the trained ANN for practical applications. Our design method can also be applied to design various nanophotonic structures that are particularly sensitive to resonant wavelengths, such as spectroscopic detection and multi-color applications. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415789>

### 1. INTRODUCTION

Metamaterial perfect absorbers (MPAs) have ultra-thin structures that can absorb almost all incident light [1,2]. This property has been exploited in several prominent applications such as thermal emitters [3], photovoltaics [4], spectroscopy [5], and sensors [6]. To achieve perfect absorption, MPAs have been developed as a variety of nano-structured devices that can control and manipulate the electromagnetic wave at the subwavelength scale. These structures are composed of either resonators [2,7] that couple to electric and magnetic fields, or of metal-insulator-metal (MIM) structures [8,9] that localize the electromagnetic fields inside the dielectric waveguide. Most proposed MPAs have one single resonant absorption peak in a single structure.

Multiband absorbers that exhibit multiple resonances in a single structure would be beneficial for spectroscopic or multi-color applications [10,11]. Therefore, approaches to integrate several resonators into a single structure have been proposed [12–14]. In these structures, each component excites a single resonance at each corresponding wavelength, and thereby yields overall multiband or broadband resonances. However, the approaches usually yield large or complex structures. In contrast,

multiband absorbers can also be achieved using a relatively simple method of stacking several sets of MIM structures [15,16]. In particular, MIMs that have the structured top layers can provide much higher absorption of light, because the top structured layers can enhance the localized field inside the waveguide layer as well as the coupling of the incident light with the guided mode. Still, designing multiband absorbers for targeting multiple wavelengths of interest using MIM is not straightforward. In addition, the design process becomes more complicated for tasks that involve multiple designs, i.e., the design process must be repeated case-by-case for target tasks. This problem also applies to cases in which multiple design tasks must be applied to design general photonic devices. Efficient and flexible design methods are being sought.

Along with the rapid development of a machine learning technology, the design problem in nanophotonics has been mitigated recently with the capability of learning complex functions from the data [17–19]. These methods introduce artificial intelligence to represent the intricate functions and, thereby, to allow non-intuitive inverse designs in nanophotonics without the need to solve computationally expensive electromagnetic problems [20–25]. However, the approach that uses deep learning generally entails high computational cost to obtain

**Table 1.** Hyperparameters Used in the Training of Two Networks

	Design Network	Spectrum Network
Number of neurons	[202, 400, 1000, 2000, 1000, 500, 200, 5]	[5, 200, 500, 1000, 500, 200, 101]
Optimizer	Adam, weight decay $10^{-5}$	Adam, weight decay $10^{-5}$
Learning rate	From $10^{-5}$ to $10^{-4}$	$10^{-4}$
Nonlinear activation functions	Leaky ReLU, $\alpha = 0.2$	Leaky ReLU, $\alpha = 0.2$

sufficient data for use in training and optimizing the network, i.e., the network itself, including the number of neurons and layers, should be optimized for the task at hand. Therefore, the overall computational cost easily exceeds that of other design methods for a single design. However, these up-front costs are incurred only once, so the method can provide subsequent on-demand designs within a few seconds.

This paper presents an efficient and spectrally sensitive design method that uses an artificial neural network (ANN) for multiband absorbers. A five-layered MIM grating structure is used for a multiband absorber, and its geometric parameters are designed using an ANN. An ANN is developed to design structures and focuses particularly on target resonant wavelengths. The developed ANN retrieves the geometric parameters from a given target input spectrum and spectral resonant wavelengths. The trained ANN is tested to evaluate its ability to design structures when given unseen optical properties, and then the designed structures are evaluated by finite difference time domain (FDTD) numerical simulation. The effective design capability of the developed network for multiple design tasks is investigated by achieving on-demand optical properties at various target wavelengths. We also show that the trained ANN can learn physical knowledge from the data by analyzing design results on gradually changing target wavelengths of the inputs. Finally, we demonstrate a method to reduce the size of an ANN through an observation for practical applications.

## 2. RESULTS AND DISCUSSIONS

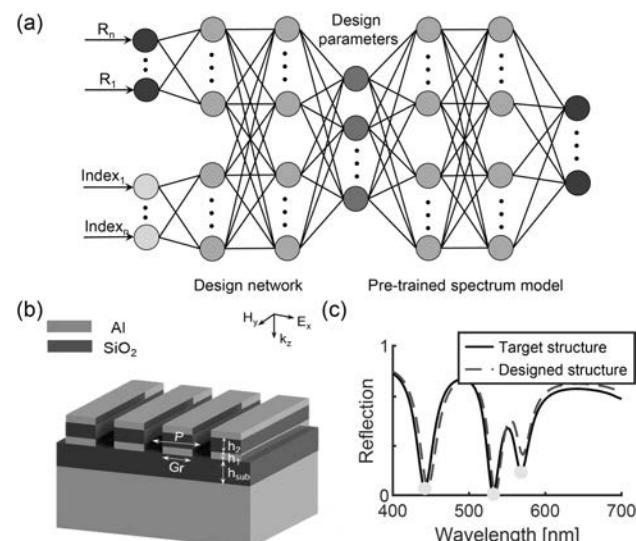
### A. Deep Learning Procedure

The total 12,100 pairs of grating structures and their associated reflection spectra were prepared using FDTD simulations for deep learning. The network to design a specified grating structure is composed of two ANNs: one is a design network, and one is a pre-trained spectrum network (Table 1). The design network retrieves design parameters of period ( $P$ ), grating width ( $G_r$ ), and each layer thickness ( $h_1$ ,  $h_2$ , and  $h_{\text{sub}}$ ) [Fig. 1(b)] from given inputs, and the pre-trained spectrum network evaluates the output of the design network by predicting optical properties of the retrieved design parameters. The input reflection spectrum with a wavelength of interest from 400 nm to 700 nm is discretized into 101 data points ( $R_1$  to  $R_n$ ). In optical design, the resonant wavelength of interest could be more important than the optical spectrum itself. Therefore, spectral information of resonant wavelengths [Index<sub>1</sub> to Index<sub>n</sub> in Fig. 1(c)] is additionally fed into the network to achieve spectrally sensitive inverse design. The information is in binary form with a total length of 101, where 1 represents the resonant wavelength. The spectrum network is well-trained in advance, showing good agreement with the FDTD simulation results

with an average mean squared error (MSE) of 0.001. The aid of the well-trained spectrum network increases the reliability and robustness of on-demand designs [21]. Overall, the two cascaded networks evaluate the discrepancies  $l_{\text{design loss}}$  between the target and designed parameters, and  $l_{\text{spectrum loss}}$  between the target and the designed spectrum [Eq. (1)]. The design loss is also added to improve the robustness of the network [26]. The network is trained to minimize the total loss, which is a weighted sum of two losses with the weight of  $w = 0.01$ . As a pre-processing of the data, the output grating parameters are divided by standard deviation of each parameter to mitigate different ranges of parameters. We adopt the batch gradient descent method with a batch size of 128, and all networks are trained on a single Nvidia GTX 1080 Ti with 11 GB memory.

$$l_{\text{total loss}} = wl_{\text{design loss}} + (1 - w)l_{\text{spectrum loss}},$$

$$\begin{aligned} l_{\text{design loss}} &= \frac{1}{n} \sum_{i=1}^{n=5} (D_{\text{target},i} - \hat{D}_{\text{designed},i})^2, \\ l_{\text{spectrum loss}} &= \frac{1}{n} \sum_{i=1}^{n=101} (R_{\text{target},i} - \hat{R}_{\text{designed},i})^2. \end{aligned} \quad (1)$$

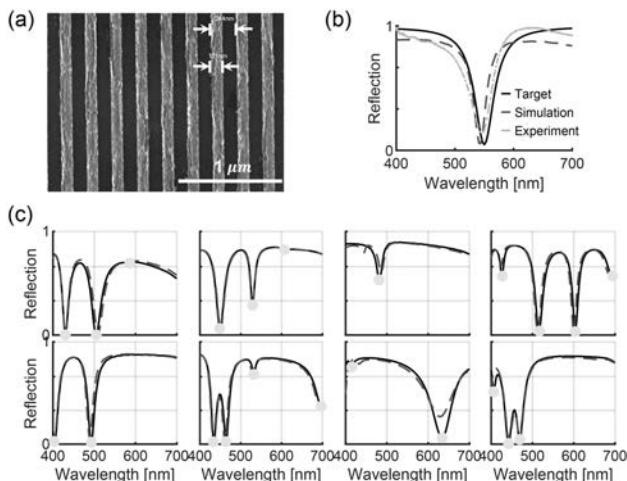


**Fig. 1.** Schematic of the designing grating structures for multiband absorbers. (a) A schematic of ANN for designing grating structures. The network is composed of two artificial neural networks of design network and pre-trained spectrum network. The design network both takes the input reflection spectra and resonant wavelengths, and the pre-trained spectrum network takes design parameters to evaluate the optical reflection spectra of the designed structures. (b) A schematic and (c) an example of optical property of a perfect multiband absorber under investigation. Yellow markers indicate resonant wavelengths.

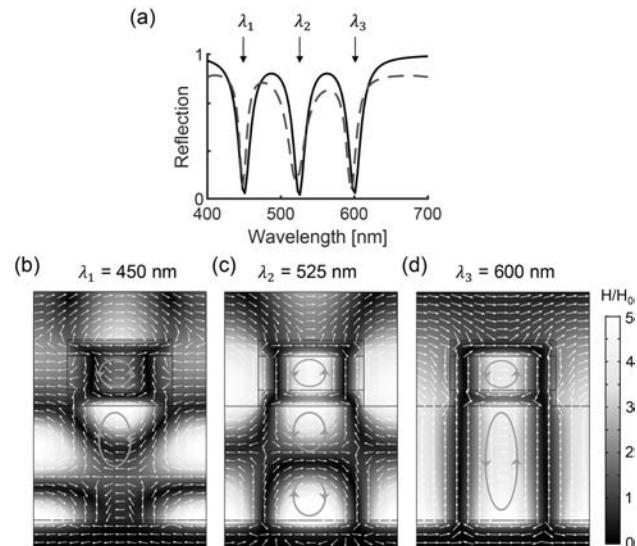
## B. Network Evaluation

The total data of 12,100 was split into 80% for training, 10% for validation, and 10% for testing. In every training iteration, the network was fitted on the training dataset, and then the fitted network was validated against the validation dataset. After training was complete, the trained network was evaluated against the unseen test dataset. For each given target optical property, the trained network provided a design; the parameters were then used in FDTD simulation to obtain electromagnetic responses. Figure 2(a) shows an example of the scanning electron microscope image of the designed grating structure. For a given target optical property, the FDTD simulation result and experimental result of the designed parameters show good agreement [Fig. 2(b)]. For a total of 1210 test data, the average MSE between target optical properties and designed responses was about 0.023, which shows that the network can well retrieve appropriate structures for given desired optical properties. Test examples in Fig. 2(c) show that the network can well design grating structures with high spectral accuracy.

The proposed grating MIM structure can be regarded as two cascaded MIM waveguides and, hence, can generate various electromagnetic resonances depending on structural parameters. For example, electric displacement in metal layers of the MIM structure can form a loop and, hence, generate magnetic resonance that amplifies strong electromagnetic absorption at resonant wavelengths. To analyze mechanisms of multiple resonances in the structure, we calculated the magnetic field distributions and electric displacements at the resonant wavelengths [Fig. 3(a)]. Figures 3(b)–3(d) clearly show that the magnetic fields are strongly localized in the waveguide region. The electric displacements in the structure (white arrows) form closed loops (green circles), generating strong magnetic

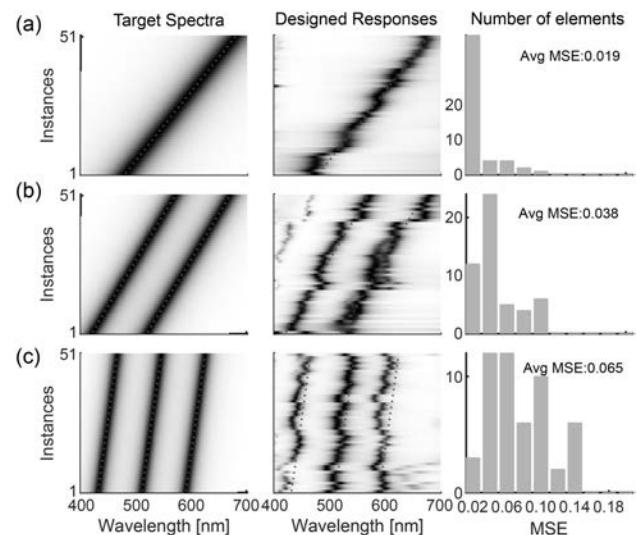


**Fig. 2.** (a) Scanning electron microscope image of a designed grating structure with a scale bar of 1  $\mu\text{m}$ . (b) Target reflection spectrum (black solid line) and designed optical properties obtained from the FDTD simulation (red dotted line) and experiment (yellow dotted line). Grating parameters with  $[P, G_r, b_1, b_2, h_{\text{sub}}] = [245 \text{ nm}, 120 \text{ nm}, 42 \text{ nm}, 113 \text{ nm}, 195 \text{ nm}]$  are designed by the network. (c) Examples of test results are shown. Black solid lines and red dotted lines are the input and target reflection spectra, respectively, and yellow markers are indexed resonant wavelengths.



**Fig. 3.** (a) Target (black solid line) and designed reflection spectrum. Magnetic field distribution (color maps) and electric displacement (arrow surfaces) at the resonant wavelengths of (b) 450 nm, (c) 525 nm, and (d) 600 nm.

resonances. At the wavelength of  $\lambda_1 = 450 \text{ nm}$ , electric displacements in two  $\text{SiO}_2$  waveguide regions are symmetric with respect to the second Al layer. On the other hand, electric displacements in two  $\text{SiO}_2$  waveguide regions are anti-symmetric with respect to the second Al layer at the wavelengths of  $\lambda_2 = 525 \text{ nm}$  and  $\lambda_3 = 600 \text{ nm}$ . At the wavelengths of  $\lambda_3 = 600 \text{ nm}$  and  $\lambda_2 = 525 \text{ nm}$ , the first and second orders of Fabry–Perot modes are generated inside the second  $\text{SiO}_2$  waveguide region. Overall, the cascaded MIM grating structures allow many interesting optical responses at

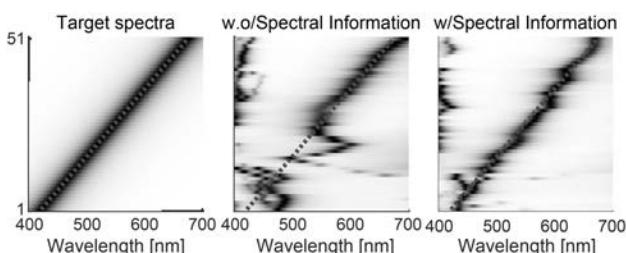


**Fig. 4.** Design of multiband absorbers with (a) single, (b) double, and (c) triple resonances. The first column shows the target input spectra, and the second column shows the designed responses. The red lines indicate target resonant wavelengths. The third column shows the histogram of the MSE for a total of 51 input spectra. The insets show the average MSE of the test input.

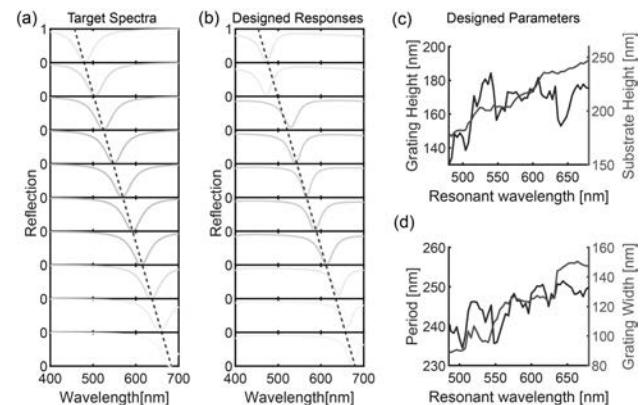
various resonant wavelengths; for this reason, the proposed structures are compatible with design by deep learning. Design approaches that use deep learning usually entail a very high initial computational cost to obtain large amounts of data, but the trained network can be used repeatedly to design various structures. Therefore, for multiple design tasks that target various resonant wavelengths, the trained network can very quickly design structural parameters that have the desired optical properties.

To demonstrate the ANN's capability for multiple design tasks, we tested our network with various target reflection spectra drawn with Lorentzian-shaped resonances (Fig. 4). We aimed to design various absorbers that had single [Fig. 4(a)], double [Fig. 4(b)], and triple [Fig. 4(c)] resonances. For each multiband absorber, 51 instances of optical properties with different resonant wavelengths were used as desired target inputs. For a total of 153 target optical properties, the designed responses show very good agreement. In particular, the optical properties are well represented in the target wavelength regions due to the additional inputs of the target resonance wavelengths. The MSE histograms (Fig. 4) show the statistical error distribution of a dataset having 10 bins with an interval of 0.02. The histogram counts the number of cases with MSE that fall into each range out of a total of 51 cases. Overall, the average MSE tended to increase as the number of resonances was increased due to the difficulty of targeting multiple resonances. Still, results clearly show that the ANN has highly generalizable capability to design various target spectra. The network performance for designing multiband absorbers could be improved if more data on multiband spectra are added to the training data.

To investigate the effect of additional input information, we also trained the network without the spectral information of resonant wavelengths (Fig. 5). The network uses all other parameters the same as those used in the main text, except that the network only takes the reflection spectrum as input. Without resonant wavelength information (the middle of Fig. 5), the predicted optical properties of the designed structures can well mimic the overall behaviors of the target optical properties, but resonant wavelengths deviate from the target wavelengths (red dotted lines) resulting in poor spectral accuracy. On the other hand, with the resonant wavelength information, the network can show much higher spectral accuracy (the right of Fig. 5).



**Fig. 5.** Comparison between two networks fed with and without spectral resonant wavelengths. The left is the target input spectra; the middle and the right are the predicted response of the networks without and with spectral information, respectively. The red lines are target resonant wavelengths.

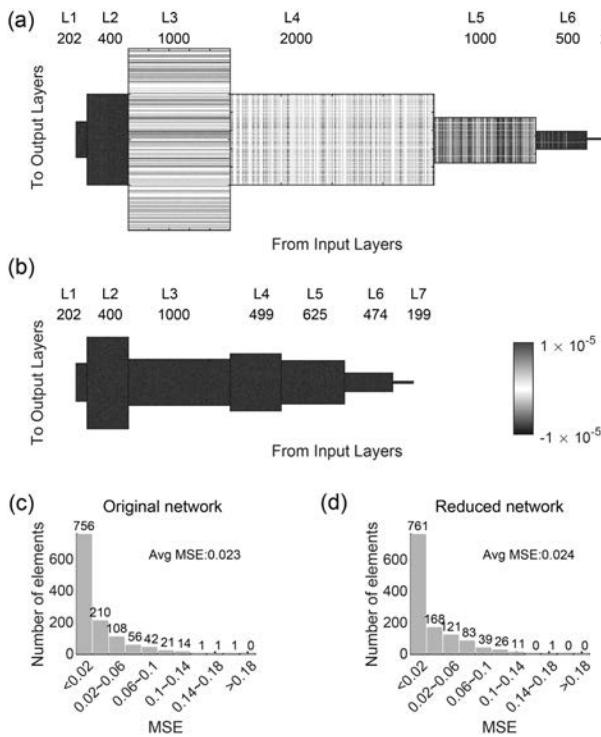


**Fig. 6.** Analysis on output parameters for gradually changing target resonant wavelengths. (a) Target spectra with gradually changing resonant target wavelengths and (b) corresponding designed responses. For given varying input spectra, the designed parameters of (c) grating height and substrate height and (d) period and grating width.

The design capability of the trained ANN was further investigated by analyzing the output parameters as the target resonant wavelengths were gradually changed (Fig. 6). For gradually changing the target resonant wavelength from 480 nm to 680 nm of the single resonant absorber [Fig. 6(a)], ANN successfully designed structures that can reproduce the desired properties [Fig. 6(b)]. Figures 6(c) and 6(d) show the designed parameters for given inputs. Interestingly, when the target resonant wavelength was redshifted, the designed substrate heights tended to increase, whereas the designed total grating heights rarely changed. In addition, the periods were almost unchanged at ~245 nm, but the grating widths increased continuously. These tendencies correspond to our physical intuitions and knowledge, i.e., it is well known that increasing the height of the insulator inside the MIM structure will lead to the redshifted electromagnetic resonances. Also, the increased grating width leads to an increased effective refractive index of the waveguide structure resulting in redshift of the resonant wavelength. Therefore, these results suggest that the ANN can learn physics by analyzing data.

### C. Network Pruning

We demonstrate an approach to prune the ANN by eliminating unnecessary neurons through observation of the trained network. It has been known that building large ANNs can generally perform better on a variety of tasks, but it is also more expensive to use. Therefore, it is important to reduce the size of networks while minimizing performance degradation. ANN pruning is a method of compressing network size by removing weights [27–29]. Here, we demonstrate pruning the ANN through observation. After the network was trained, we visualized the trained values of the weights connecting neurons in the previous input layer ( $x$  axis) to neurons in the output layer ( $y$  axis) (Fig. 7). Interestingly, we observed that some weights linking between two neurons are almost zero [white in Fig. 7(a)]. For example, all weights connecting from all neurons in layer 2 (L2) to the first neuron in layer 3 (L3) are zero (the first white horizontal line in L3), so the neuron is turned off.

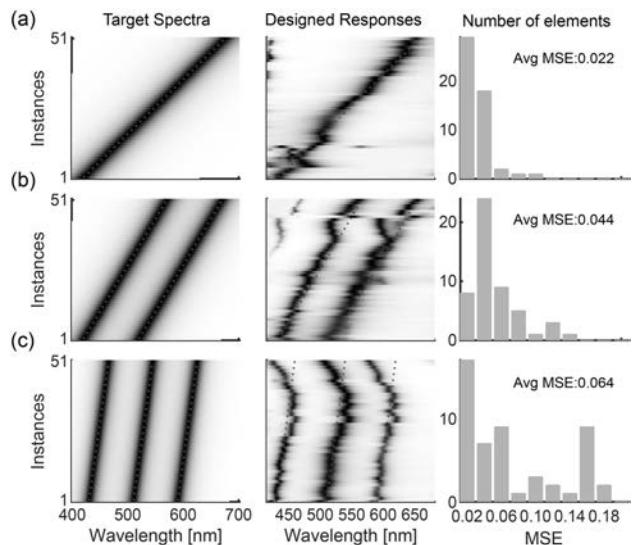


**Fig. 7.** Network pruning results. Visualization of the trained weights in (a) the original network and (b) the pruned network. For each layer ( $L_n, n = 1, 2, \dots, 7$ ), the number of neurons is indicated. MSE histogram of the test data for (c) the original network and (d) the pruned network.

It means that the neuron did not receive any information from all neurons in the previous layer and, consequently, did not transmit any information to the next layer (the first white vertical line in L4). This implies that the neuron did not participate in learning, so layer 3 has more neurons than necessary. Therefore, we counted the number of such turned-off neurons in each layer and removed the neurons to reduce the total size of the network. By reducing the size of the network, the number of trainable parameters is reduced from about 5 million to 1.6 million (Table 2). Not surprisingly, despite the significant network compression, the pruned network retains the design capability similar to the original network in the MSE histogram of the test data [Figs. 7(c) and 7(d)]. In addition, the pruned network also enables efficient multiple design tasks for various user-drawn spectra (Fig. 8). These results suggest that the network can be pruned through observation of the trained weights. The proposed network method could be applied to any other research fields that use ANN.

**Table 2. Number of Neurons in Each Layer**

Layer	1	2	3	4	5	6	7	Total
Original	202	400	1000	2000	1000	500	200	5055905
Reduced	202	400	1000	499	625	474	199	1651642



**Fig. 8.** Design of multiband absorbers with (a) single, (b) double, and (c) triple resonances using the reduced network. The first column shows target input spectra, and the second column shows the designed response. The red lines indicate target resonant wavelengths. The third column shows the histogram of the MSE for a total of 51 input spectra.

### 3. CONCLUSION

We have proposed a method that uses deep learning to design spectrally sensitive multiband absorbers. By feeding additional spectral information of resonant wavelengths, the developed ANN achieved a highly robust and accurate design of multiband absorbers. We have also analyzed the results of the designed outputs when the resonant wavelength was gradually changed. For gradually redshifting inputs, the designed parameters of the substrate height and period tended to increase, in correspondence to physical intuition. The results suggest that the trained network can well grasp and learn physics by analyzing data. We envision that this can also be applied to other nanophotonic problems and may solve complex light–matter interactions that are even beyond our knowledge. Finally, we proposed a systematic method that uses observation of the trained network to guide its pruning. The method is expected to significantly reduce the computational cost involved in network reduction. The method could be extended to other research fields that use ANNs. In this study, we considered designing one structure for one given spectrum, but it is worth noting that multiple candidates could be designed by using deep learning algorithms for multi-output regression. We believe considering several design candidates would be beneficial for fabrication when the target design becomes much more complex.

**Funding.** National Research Foundation of Korea (NRF-2018M3D1A1058998, NRF-2019R1A2C3003129, CAMM-2019M3A6B3030637, NRF-2019R1A5A8080290, NRF-2020K1A3A1A21024374); Ministry of Education (NRF-2017H1A2A1043322, NRF-2019H1A2A1076295).

**Acknowledgment.** Y. Y. acknowledges a fellowship from the Hyundai Motor Chung Mong-Koo Foundation.

**Disclosures.** The authors declare that there are no conflicts of interest related to this paper.

## REFERENCES

1. N. Engheta, "Thin absorbing screens using metamaterial surfaces," in *IEEE Antennas and Propagation Society International Symposium* (IEEE, 2002), pp. 392–395.
2. N. I. Landy, S. Sajuyigbe, J. J. Mock, D. R. Smith, and W. J. Padilla, "Perfect metamaterial absorber," *Phys. Rev. Lett.* **100**, 207402 (2008).
3. M. Diem, T. Koschny, and C. M. Soukoulis, "Wide-angle perfect absorber/thermal emitter in the terahertz regime," *Phys. Rev. B* **79**, 033101 (2009).
4. A. Vora, J. Gwamuri, N. Pala, A. Kulkarni, J. M. Pearce, and D. Ö. Güney, "Exchanging ohmic losses in metamaterial absorbers with useful optical absorption for photovoltaics," *Sci. Rep.* **4**, 4901 (2014).
5. F. B. Barho, F. Gonzalez-Posada, M. Bomers, A. Mezy, L. Cerutti, and T. Taliercio, "Surface-enhanced thermal emission spectroscopy with perfect absorber metasurfaces," *ACS Photon.* **6**, 1506–1514 (2019).
6. N. Liu, M. Mesch, T. Weiss, M. Hentschel, and H. Giessen, "Infrared perfect absorber and its application as plasmonic sensor," *Nano Lett.* **10**, 2342–2348 (2010).
7. H. Tao, C. Bingham, D. Pilon, K. Fan, A. Strikwerda, D. Shrekenhamer, W. Padilla, X. Zhang, and R. Averitt, "A dual band terahertz metamaterial absorber," *J. Phys. D* **43**, 225102 (2010).
8. J. Hao, J. Wang, X. Liu, W. J. Padilla, L. Zhou, and M. Qiu, "High performance optical absorber based on a plasmonic metamaterial," *Appl. Phys. Lett.* **96**, 251104 (2010).
9. P. Bouchon, C. Koechlin, F. Pardo, R. Hadar, and J.-L. Pelouard, "Wideband omnidirectional infrared absorber with a patchwork of plasmonic nanoantennas," *Opt. Lett.* **37**, 1038–1040 (2012).
10. K. Chen, R. Adato, and H. Altug, "Dual-band perfect absorber for multispectral plasmon-enhanced infrared spectroscopy," *ACS Nano* **6**, 7998–8006 (2012).
11. J. Xu, Z. Zhao, H. Yu, L. Yang, P. Gou, J. Cao, Y. Zou, J. Qian, T. Shi, Q. Ren, and Z. An, "Design of triple-band metamaterial absorbers with refractive index sensitivity at infrared frequencies," *Opt. Express* **24**, 25742–25751 (2016).
12. J. W. Park, P. Van Tuong, J. Y. Rhee, K. W. Kim, W. H. Jang, E. H. Choi, L. Y. Chen, and Y. Lee, "Multi-band metamaterial absorber based on the arrangement of donut-type resonators," *Opt. Express* **21**, 9691–9702 (2013).
13. Y. Ma, Q. Chen, J. Grant, S. C. Saha, A. Khalid, and D. R. Cumming, "A terahertz polarization insensitive dual band metamaterial absorber," *Opt. Lett.* **36**, 945–947 (2011).
14. S. Liu, H. Chen, and T. J. Cui, "A broadband terahertz absorber using multi-layer stacked bars," *Appl. Phys. Lett.* **106**, 151601 (2015).
15. Z. Zhang, Z. Yu, Y. Liang, and T. Xu, "Dual-band nearly perfect absorber at visible frequencies," *Opt. Mater. Express* **8**, 463–468 (2018).
16. G. Dayal and S. A. Ramakrishna, "Design of multi-band metamaterial perfect absorbers with stacked metal-dielectric disks," *J. Opt.* **15**, 055106 (2013).
17. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2021).
18. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**, 339–366 (2019).
19. S. So, T. Badloe, J. Noh, J. Rho, and J. Bravo-Abad, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **9**, 1041–1057 (2020).
20. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
21. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
22. S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core-shell nanoparticles," *ACS Appl. Mater. Interfaces* **11**, 24264–24268 (2019).
23. Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
24. S. So and J. Rho, "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**, 1255–1261 (2019).
25. Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, "Knowledge discovery in nanophotonics using geometric deep learning," *Adv. Intell. Syst.* **2**, 1900132 (2020).
26. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
27. J. Frankle and M. Carbin, "The lottery ticket hypothesis: finding sparse, trainable neural networks," arXiv:1803.03635 (2018).
28. D. Blalock, J. J. G. Ortiz, J. Frankle, and J. Guttag, "What is the state of neural network pruning?" arXiv:2003.03033 (2020).
29. M. Babaeizadeh, P. Smaragdis, and R. H. Campbell, "Noiseout: a simple way to prune neural networks," arXiv:1611.06211 (2016).

# PHOTONICS Research

## Intelligent coding metasurface holograms by physics-assisted unsupervised generative adversarial network

CHE LIU,<sup>1,2</sup>  WEN MING YU,<sup>1,2</sup> QIAN MA,<sup>1,2</sup> LIANLIN LI,<sup>3</sup> AND TIE JUN CUI<sup>1,2,\*</sup>

<sup>1</sup>Institute of Electromagnetic Space, Southeast University, Nanjing 210096, China

<sup>2</sup>State Key Laboratory of Millimeter Wave, Southeast University, Nanjing 210096, China

<sup>3</sup>School of Electronic Engineering and Computer Sciences, Peking University, Beijing 100871, China

\*Corresponding author: tjcui@seu.edu.cn

Received 30 November 2020; revised 4 February 2021; accepted 6 February 2021; posted 8 February 2021 (Doc. ID 416287); published 31 March 2021

Intelligent coding metasurface is a kind of information-carrying metasurface that can manipulate electromagnetic waves and associate digital information simultaneously in a smart way. One of its widely explored applications is to develop advanced schemes of dynamic holographic imaging. By now, the controlling coding sequences of the metasurface are usually designed by performing iterative approaches, including the Gerchberg–Saxton (GS) algorithm and stochastic optimization algorithm, which set a large barrier on the deployment of the intelligent coding metasurface in many practical scenarios with strong demands on high efficiency and capability. Here, we propose an efficient non-iterative algorithm for designing intelligent coding metasurface holograms in the context of unsupervised conditional generative adversarial networks (cGANs), which is referred to as physics-driven variational auto-encoder (VAE) cGAN (VAE-cGAN). Sharply different from the conventional cGAN with a harsh requirement on a large amount of manual-marked training data, the proposed VAE-cGAN behaves in a physics-driving way and thus can fundamentally remove the difficulties in the conventional cGAN. Specifically, the physical operation mechanism between the electric-field distribution and metasurface is introduced to model the VAE decoding module of the developed VAE-cGAN. Selected simulation and experimental results have been provided to demonstrate the state-of-the-art reliability and high efficiency of our VAE-cGAN. It could be faithfully expected that smart holograms could be developed by deploying our VAE-cGAN on neural network chips, finding more valuable applications in communication, microscopy, and so on.

© 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416287>

### 1. INTRODUCTION

Electromagnetic (EM) metasurfaces use periodic or quasi-periodic macroscopic basic units to simulate the atoms or molecules on the microscale of traditional material science. These macroscopic basic units could interact with external electric fields via resonance effects and express unique EM properties [1–6] to realize various novel functional devices such as cloaks [7–10], concentrators [11], illusion optics devices [12,13], special lenses [14,15], and diffuse reflections [16]. Recently, a series of special EM metasurfaces called coding, digital, and programmable metasurfaces [17] ('coding metasurfaces' for short in the remainder of this article) have gained more attention due to the unique methods for manipulating the EM waves. By encoding the phase responses of the metasurfaces as digital numbers '0' or '1', the EM property of each meta-unit in the coding metasurfaces could be switched in real time when controlled by a field-programmable gate array (FPGA), which allows people to

design metasurfaces in digital space rather than the analog domain. Further, the digital representation of meta-units could link the EM space and digital world and has been widely used in manipulating amplitude [18,19], polarization [20–22], and orbital angular momentum [23–25], yielding the concept of information metasurfaces [26–29] and many other novel applications [30–33].

With the development of communication technology, people need to deal with more and more information in daily life, which adds the burden of the whole society and promotes the development of artificial intelligence (AI) to help people deal with various information handling tasks such as speech recognition [34–36], image recognition [37–39], automatic translation [40–42], and robot control [43–45]. The growth of AI has also brought changes to the design of EM metasurfaces, especially coding metasurfaces. The digital representation of the coding metasurfaces makes it more convenient to put AI technology into the design of the meta-unit states [46–48] and has

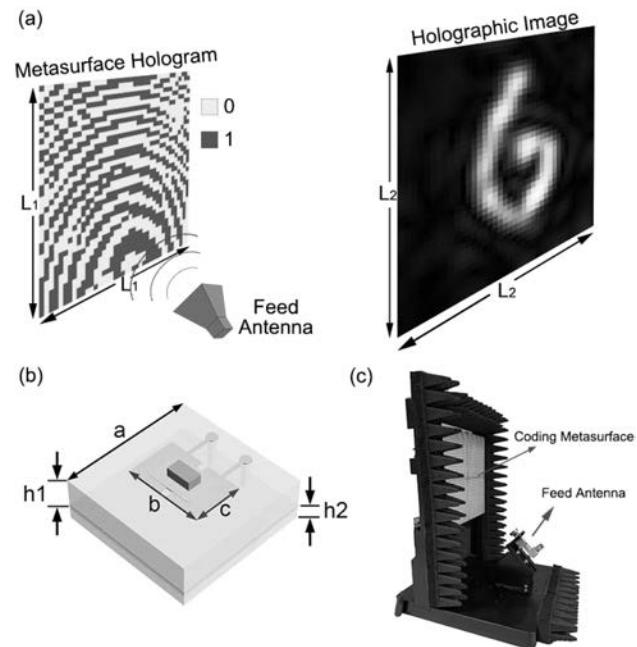
already sprung out various interesting applications such as the smart system [49], high-resolution imager [47,50], and recognizer [51].

Despite the wide usage of AI in coding metasurfaces, the design of meta-unit states for holographic imaging by coding metasurfaces still rests on iteration optimization algorithms like the Gerchberg–Saxton (GS) algorithm [52,53] or greedy algorithm [54]. Recently, a valuable trail has been made for using supervised deep neural networks (DNNs) to recover holograms from the generated speckles distorted by a thin diffuser [55]. However, the demand for thousands of training samples acquired by real measurements remarkably increases the cost in the usage of this supervised DNN. Here, we propose a new method based on AI or deep learning that could rapidly generate the coding pattern (metasurface hologram) of the 1 bit coding metasurface when a target holographic image is given. The deep learning structure composed of unsupervised variational auto-encoder (VAE) and conditional generative adversarial networks (cGANs) is presented in our method. The use of VAE [56], as an unsupervised algorithm, could remarkably reduce the time consumption for the preparation of training data, and the EM propagation model described by the rigorous dyadic Green's function (DGF) [53] is used to make the VAE structure possible while keeping the physical interpretability. We merge the Wasserstein distance [57–59] [by the form of Wasserstein GAN (WGAN) [60,61]] and mean square error (MSE) in the design of loss function to make the distributions of generated holographic images and target holographic images closer. The structure of cGAN [62–64] is also used to avoid the confusion of the generative network. We design a  $40 \times 40$  one bit coding metasurface working at the frequency of 35 GHz to validate our intelligent method, and the simulation and experimental results both validate the efficiency and reliability of the proposed approach.

## 2. THEORY AND METHOD

### A. EM Propagation Model

Figure 1(a) shows the sketch of our hologram system, which consists of a  $40 \times 40$  one bit coding metasurface controlled by FPGA and a feed antenna radiating the EM waves in the frequency range from 34 to 36 GHz. Each meta-unit on the coding metasurface could be switched between two opposite states, radiating the EM waves with opposite phases of 0 or  $\pi$ , which are represented by codes '0' and '1', respectively. The 1 bit coding metasurface is the imager that loads the metasurface hologram, which can project the holographic image at the imaging plane 30 cm away from it. The size of the metasurface hologram and imaging plane is  $L_1 = 152$  mm and  $L_2 = 200$  mm. Figure 1(b) shows the sketch of the meta-unit's structure, whose parameters are given by  $a = 3.8$  mm,  $b = 2.2$  mm,  $c = 1.4$  mm,  $h_1 = 0.813$  mm, and  $h_2 = 0.435$  mm. The substrate material of the metasurface is Rogers RO 4003C ( $\epsilon_r = 3.55$ ,  $\tan \delta = 0.0027$ ). A PIN diode is integrated on the top layer of the metasurface element, connecting with two metal patches, which offer the bias voltage control through two via-holes. When the on/off state of diode is changed by specific bias voltage, the surface electric field alters distinctively to generate the reflected phase response with  $180^\circ$  difference.



**Fig. 1.** (a) Sketch of our hologram system that consists of a 1 bit coding metasurface loading the hologram and a feed antenna. The distance between the holographic imaging plane and coding metasurface is 30 cm. (b) The meta-unit of the 1 bit coding metasurface. (c) The side-looking photograph of our hologram system. The interval between meta-units on the coding metasurface is 3.8 mm, and the feed antenna radiates the EM waves with frequencies from 34 GHz to 36 GHz.

The photograph of the hologram system is illustrated in Fig. 1(c), which is a well-integrated system with the coding metasurface controlled by the FPGA via an ethernet communication interface.

We accept rigorous DGF [53] as the basic computational kernel of the forward propagation from the source currents of meta-units to the near-field EM distribution:

$$\vec{G}(\mathbf{r}, \mathbf{r}') = \left( \vec{\mathbf{I}} + \frac{\nabla \nabla}{k^2} \right) g(\mathbf{r}, \mathbf{r}'), \quad (1)$$

where  $\vec{\mathbf{I}}$  is a  $3 \times 3$  dyadic identity matrix,  $\mathbf{r}$  and  $\mathbf{r}'$  are source and field points, respectively, and

$$g(\mathbf{r}, \mathbf{r}') = \frac{e^{-jkR}}{4\pi R} \quad (2)$$

is the free-space Green's function, in which  $R = |\mathbf{r} - \mathbf{r}'|$ . The forward propagation formula could be represented as

$$\mathbf{E}(\mathbf{r}) = -j\omega\mu \int_V d\mathbf{r}' \vec{G}(\mathbf{r}, \mathbf{r}') \cdot \mathbf{J}(\mathbf{r}'), \quad (3)$$

where  $\mathbf{E}(\mathbf{r})$  and  $\mathbf{J}(\mathbf{r}')$  represent the electric field at the field point and the current at the source point, respectively. Owing to the discrete array form of the coding meta-units, the source currents can also be expressed in a discrete form. For convenience, we use a current element to represent a coding meta-unit and discrete the EM field into  $M$  points at the same time. Then, Eq. (3) could reduce to

$$E(\mathbf{r}_m) = -j\omega\mu \sum_{n=1}^N \vec{G}(\mathbf{r}_m, \mathbf{r}'_n) \cdot J(\mathbf{r}'_n), \quad m = 1, \dots, M, \quad (4)$$

where  $N$  is the number of coding meta-units. Further, we organize the scalar components of  $E(r_m)$  and  $J(r'_n)$  in Eq. (4) for all  $m$  and  $n$  into column vectors  $\mathbf{E}$  and  $\mathbf{J}$ , respectively. As we can see from Eq. (4), the relationship between  $E(r_m)$  and  $J(r'_n)$  is linear, and, hence, they could be connected with a complex-value coefficient matrix. Then, Eq. (4) could ultimately reduce to

$$\mathbf{E} = \mathbf{W} \cdot \mathbf{J}, \quad (5)$$

where  $\mathbf{W}$  is the coefficient matrix that links the source  $\mathbf{J}$  and field  $\mathbf{E}$ . We only care about the vertical polarization EM waves and the relative amplitude of the near-field EM distribution. Hence, the current vector  $\mathbf{J}$  could be described as

$$\mathbf{J} = \mathbf{J}' \odot \mathbf{J}^\varphi, \quad (6)$$

where  $\mathbf{J}'$  is a complex-value vector and represents the current part that is directly caused by the incident EM waves, which is proportional to the incident electric-field values;  $\odot$  means element-wise multiplication; and  $\mathbf{J}^\varphi$  represents the current part that is controlled by each coding meta-unit, which is a real-value vector whose elements are '1' or '-1' to represent the phase of 0 or  $\pi$ , corresponding to code '0' or '1'. Then, Eq. (5) could be rewritten as

$$\begin{aligned} \mathbf{E} &= \mathbf{W} \cdot \mathbf{J}' \odot \mathbf{J}^\varphi = \mathbf{W} \cdot \text{diag}(\mathbf{J}') \cdot \mathbf{J}^\varphi = \mathbf{W}' \cdot \mathbf{J}^\varphi, \\ \mathbf{W}' &= \mathbf{W} \cdot \text{diag}(\mathbf{J}'). \end{aligned} \quad (7)$$

Thus, the forward propagation process could be represented as a form of matrix multiplication from a real-value vector  $\mathbf{J}^\varphi$  to the near-field electric-field distribution  $\mathbf{E}$ . The design objective of VAE-cGAN is to get the current vector  $\mathbf{J}^\varphi$  to generate a target near-field electric-field distribution  $\mathbf{E}$ , whose amplitude distribution represents a target holographic image.

## B. Network Structure

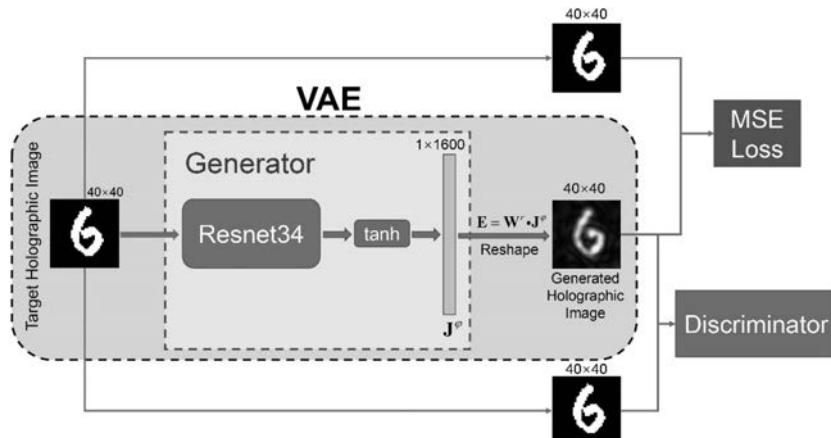
Figure 2 demonstrates the network structure of the proposed VAE-cGAN. It mainly consists of two modules: generator and discriminator. The generator is responsible for generating the current vector  $\mathbf{J}^\varphi$  when a target holographic image is given.

The binarization target holographic image is first input to the Resnet34 [65] network of the generator, whose output is activated by the hyperbolic tangent function and finally becomes a  $1 \times 1600$  current vector  $\mathbf{J}^\varphi$ , whose element value is near '1' or '-1', representing the code '0' or '1' in a  $40 \times 40$  one bit coding metasurface. We replace the batch normalization [66] with the instance normalization [67] in the Resnet34 network to prevent interaction among batches. Then, the holographic image generated by  $\mathbf{J}^\varphi$  is calculated by Eq. (7) and divided by its maximum value to keep its value in the range [0,1]; next, it is sent into the discriminator to calculate the Wasserstein distance [57] between itself and the target holographic image. The MSE is also calculated at the same time as another index, indicating the similarity between the generated holographic image and target holographic image. So, the generator together with the EM forward propagation process makes up the structure of the VAE [55], whose optimization target is making the output image as similar as possible to the input image. The usage of the unsupervised structure VAE gets us out of the trouble of making the marked training dataset, whose data must be organized in pairs with the target holographic image and its corresponding coding pattern (metasurface hologram) generated by iterative algorithms like the GS algorithm. Besides, the multiplicity-solution property of the iterative algorithm will cause one-to-many problems in the marked training dataset and makes the training process difficult to converge.

The discriminator adopts the structure of cGAN [63] and WGAN [61], which is responsible for calculating the approximate Wasserstein distance between the generated holographic image and the target holographic image. It will be trained in the adversarial process towards the generator and improve the imaging quality of the generated holographic image.

## C. Backpropagation Process

The mainstream deep learning platform such as TensorFlow and Pytorch could not directly deal with the complex values, which will cause difficulty in the backpropagation process from loss function to  $\mathbf{J}^\varphi$ . Thus, we need to derive the backpropagation partial derivative equation from the generated holographic image (the amplitude of  $\mathbf{E}$ ) to  $\mathbf{J}^\varphi$  to get rid of the calculation of complex values, which is expressed as (using numerator layout)



**Fig. 2.** Schematic diagram of the proposed VAE-cWGAN. The generator together with the EM propagation process makes up the VAE structure. Two kinds of distance criteria (MSE and Wasserstein distance) are used to improve the imaging quality of the generator.

$$\frac{\partial |\mathbf{E}|}{\partial \mathbf{J}^\varphi} = |\mathbf{E}|^{\odot -1} \odot \text{Re} \left[ \mathbf{E}^* \odot \frac{\partial \mathbf{E}}{\partial \mathbf{J}^\varphi} \right] = |\mathbf{E}|^{\odot -1} \odot \text{Re} [\mathbf{E}^* \odot \mathbf{W}^*], \quad (8)$$

where the superscript notation  $\odot -1$  means taking the reciprocal element-wisely and  $*$  means taking the conjugate matrix. Then, the backpropagation process from loss function to  $\mathbf{J}^\varphi$  can be calculated by real values,

$$\frac{\partial \text{Loss}}{\partial \mathbf{J}^\varphi} = \frac{\partial \text{Loss}}{\partial |\mathbf{E}|} \cdot \frac{\partial |\mathbf{E}|}{\partial \mathbf{J}^\varphi}. \quad (9)$$

#### D. Design of Loss Function

We merge the MSE loss and Wasserstein distance as our final loss function. The MSE loss and Wasserstein distance are both criterions to indicate the distance between two distributions. Although using the Wasserstein distance evaluation function alone could also reach an equally good result, we still add the MSE evaluation to the loss function of the generator because the MSE evaluation could act as the ‘lubricant’ for the training process to accelerate and stabilize the convergence of the generator and discriminator. The discriminator cannot tell fake samples apart from real ones at the beginning of the training process; thus, it should be trained to learn a K-Lipschitz continuous function [60,61] so as to compute the Wasserstein distance. Therefore, at the beginning of training process, the generator is not able to get any effective guides from the discriminator for the updates of parameters, which would raise the risk of ‘mode collapse’. Luckily, the MSE evaluation function could help to guide the updates of the generator before the discriminator is well trained so as to prevent the generator from the ‘mode collapse’. This is the reason why we add it into the final loss function of the generator.

The MSE loss is described as

$$\begin{aligned} \text{MSE} &= \frac{1}{N} \sum [(|\mathbf{E}| - |\mathbf{E}_t|)^{\odot 2}], \\ \frac{\partial \text{MSE}}{\partial |\mathbf{E}|} &= \frac{2}{N} [(\mathbf{E} - \mathbf{E}_t)]^T, \end{aligned} \quad (10)$$

where  $|\mathbf{E}_t|$  represents the target holographic image and superscript  $\odot 2$  means element-wise square.

Wasserstein distance [59] indicates the minimum movement for changing one distribution  $P_1$  to another distribution  $P_2$ :

$$W(P_1, P_2) = \inf_{\gamma \sim \Pi(P_1, P_2)} \mathbb{E}_{(x,y) \sim \gamma} (\|x - y\|_p), \quad (11)$$

where  $\Pi(P_1, P_2)$  represents the set of all of the possible joint distributions between  $P_1$  and  $P_2$ , and notation  $\|\cdot\|_p$  means  $p$ -norm. The Wasserstein distance is a more appropriate criterion to indicate the distance between two distributions. Figure 3(a) shows the three sets consisting of discrete sequences sampled from distributions  $P_1$ ,  $P_2$ , and  $P_3$ , respectively, defined by

$$\forall \text{ sequence } \in P_1, \quad \text{sequence} = [x, 0, 0, 0, 0, 0, 0, 0]$$

$$\text{and } x \sim U(0, 1),$$

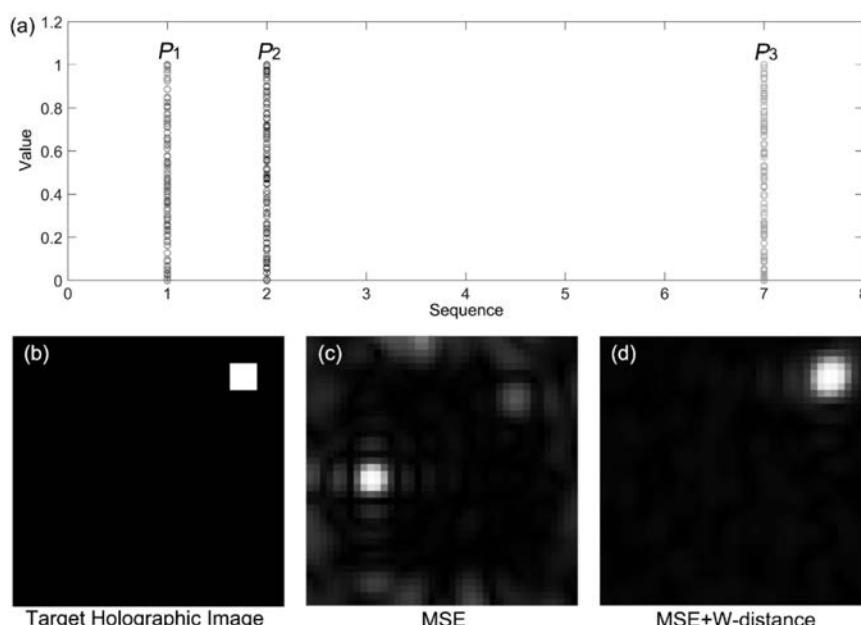
$$\forall \text{ sequence } \in P_2, \quad \text{sequence} = [0, x, 0, 0, 0, 0, 0, 0]$$

$$\text{and } x \sim U(0, 1),$$

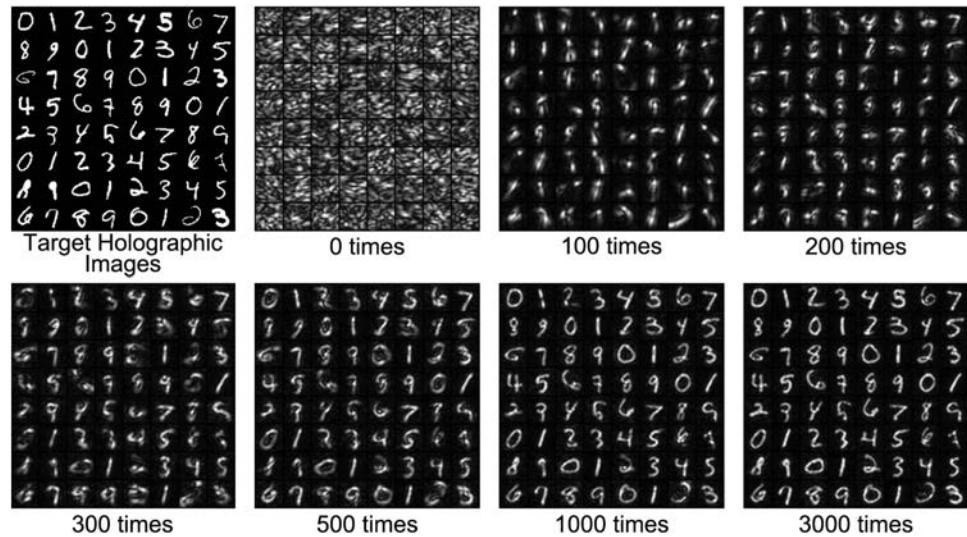
$$\forall \text{ sequence } \in P_3, \quad \text{sequence} = [0, 0, 0, 0, 0, 0, x, 0]$$

$$\text{and } x \sim U(0, 1).$$

The mathematical expectation of  $p$ -norm distance between the discrete sequences sampled from distributions  $P_1$  and  $P_2$  is the same as that between  $P_1$  and  $P_3$ . However, it is obvious that



**Fig. 3.** (a) Three discrete-sequence sets sampled from distributions  $P_1$ ,  $P_2$ , and  $P_3$ , respectively. (b) The target holographic image is input to the trained generator. (c) The generated holographic image [corresponding to Fig. 3(b)] output by a generator trained only using the MSE loss. (d) The generated holographic image output by a generator trained using MSE loss and Wasserstein distance simultaneously.



**Fig. 4.** Generated holographic images at each training time corresponding to the valid target holographic images. One time of training is when training generator has three iterations after the training discriminator has one iteration.

$P_2$  is more visually similar to  $P_1$  than  $P_3$ . Thus, the  $p$ -norm criterion may fail to indicate the distance between two sparse distributions, just as demonstrated in Figs. 3(b) and 3(c), in which the target holographic image [Fig. 3(b)] can be expressed as a sparse matrix, making it difficult to find the direction of optimization and fall into local minimum with the generator trained using only MSE loss, and eventually output a wrong holographic image [Fig. 3(c)]. Luckily, the Wasserstein distance

could help denote this distribution difference, and the generator trained by the Wasserstein distance could generate the correct holographic image [Fig. 3(d)].

For the loss function of the discriminator, we use the loss function of WGAN [61] to simulate the calculation of Wasserstein distance and introduce the concept of cGAN [63] to match the target holographic image with the generated holographic image:



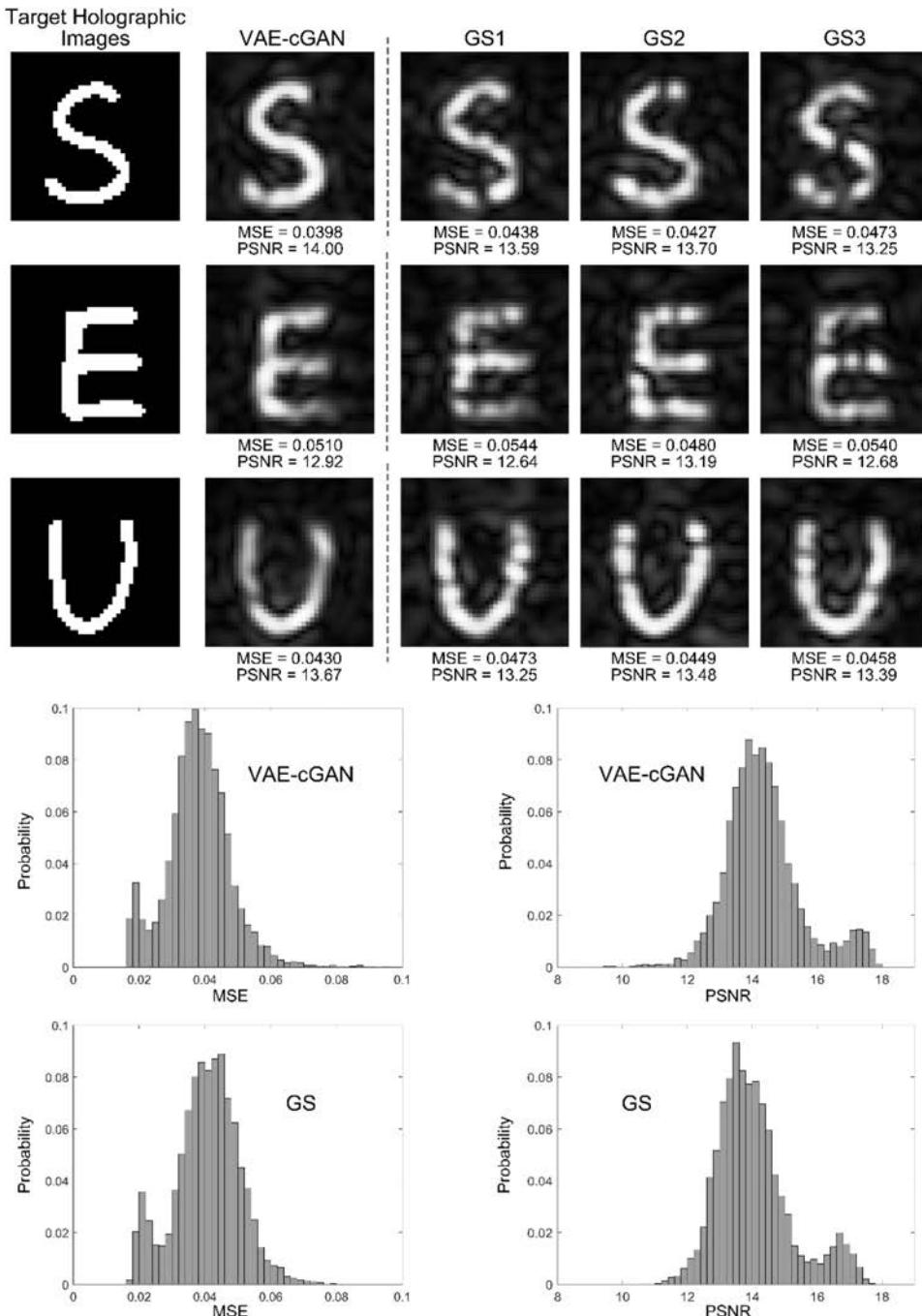
**Fig. 5.** Testing results of our proposed intelligent metasurface hologram system. The target holographic images are randomly chosen from the testing MNIST dataset or images of handwritten letters. The simulation holographic images are calculated by Eq. (7) with the binarized current vectors  $\mathbf{J}^o$  output by the generator. The experimental holographic images are radiated by our 1 bit coding metasurface configured with the corresponding metasurface holograms generated by the binarized current vectors  $\mathbf{J}^o$  from the generator.

$$\text{Loss}_D = \mathbb{E}_{\tilde{x} \sim P_g | P_r} [D(\tilde{x})] - \mathbb{E}_{x \sim P_r | P_r} [D(x)] + \lambda \mathbb{E}_{\hat{x} \sim P_{\hat{x}}} [(\|\nabla_{\hat{x}} D(\hat{x})\|_2 - 1)^2], \quad (12)$$

where  $P_g$  and  $P_r$  represent the distributions of the generated holographic images and target holographic images, respectively. Meanwhile, by adding the MSE loss in Eq. (10), our loss function for the generator could be expressed as

$$\text{Loss}_G = -\mathbb{E}_{\tilde{x} \sim P_g | P_r} [D(\tilde{x})] + \text{MSE}. \quad (13)$$

Figure 3(d) shows the generated holographic image output by the joint optimization of MSE loss and Wasserstein distance, which matches well with the target holographic image [Fig. 3(b)] and demonstrates the necessity of the mixed distance criterion.



**Fig. 6.** Comparison results between the VAE-cGAN and GS algorithms. The top half presents the simulated results of the generated holographic images radiated by metasurface holograms designed by our VAE-cGAN and GS algorithms, respectively. Here, GS1, GS2, and GS3 are the generated holographic images obtained by running the GS algorithm three times in sequence. The values of MSE and PSNR evaluations are marked below the corresponding holographic images. The bottom half of this figure illustrates the statistical frequency histograms of the holographic image quality in terms of MSE and PSNR.

### 3. RESULTS AND DISCUSSION

We use the images from the Modified National Institute of Standards and Technology (MNIST) data sets of handwritten digits [68] as the target holographic images to train our VAE-cGAN. In accordance with tradition, we divide the images of the MNIST database into three parts: training, validating, and testing datasets, with 35,000, 5000, and 10,000 samples respectively. The images from the MNIST database are reshaped to  $40 \times 40$  and binarized before being input to the generator or discriminator. Because the discriminator is much more easily trained than the generator, we train the generator three times before training the discriminator. The Adam optimizer is used with an initial learning rate of  $3 \times 10^{-4}$ , and the size of batch is 64. We use the Pytorch deep learning platform to construct our VAE-cGAN and run the program on a graphics processing unit (GPU) (Nvidia Quadro RTX 8000). The generator converges and becomes stable after about 9000 training iterations, and the total time consumption is about 2 h. The generated holographic images at each training time when giving the same 64 valid target holographic images are shown in Fig. 4, demonstrating that the generated holographic images successfully converge to the target holographic images when the training process goes on.

For the experimental process, we input a testing target holographic image into the trained generator and get its output current vector  $\mathbf{J}^\varphi$ . Then, the generated  $\mathbf{J}^\varphi$  is binarized to '1' or '-1', which corresponds to the reflection phase of 0 or  $\pi$  (the coding meta-unit state of 0 or 1), respectively. The simulated holographic images are calculated by Eq. (7) with binarized current vectors  $\mathbf{J}^\varphi$ . The coding metasurface [Fig. 1(b)] is set using an FPGA with the coding meta-unit states generated from  $\mathbf{J}^\varphi$ . Last, we measure the holographic images radiated by the digital coding metasurface holograms in a standard microwave chamber.

Figure 5 shows some of the measured results. The target holographic images of handwritten digits '0', '3', and '7' in Fig. 5 are randomly chosen from the testing MNIST dataset, which means that they have not participated in the training process of the generator. The high imaging quality of the simulation and experimental holographic images indicates that the generator possesses the abilities of generalization. Furthermore, we use extra testing holographic images of handwritten letters, which have different features from the handwritten digits, to further verify the generalization ability of the generator. The randomly selected testing results corresponding to the target holographic images of handwritten letters 'A' and 'B' are also presented in Fig. 5, indicating that the generator could handle well these handwritten letters and proving the generality of our intelligent metasurface hologram system.

The comparison results between VAE-cGAN and GS algorithms are provided in Fig. 6, in which the top half presents the simulated results of the generated holographic images radiated by the metasurface holograms designed by VAE-cGAN and GS algorithms, respectively. Since the results of the GS algorithm are inconsistent due to its random parameters initialization, for one target holographic image, we run the GS algorithm three times and get three holographic images indicated by GS1, GS2, and GS3, respectively. The values of MSE and peak signal-to-noise ratio (PSNR) evaluation are marked below the

corresponding holographic images, showing that the generated holographic images from our VAE-cGAN generally have smaller MSE and larger PSNR evaluation values than those from the GS algorithm. The bottom half of Fig. 6 illustrates the statistical frequency histograms of the holographic image quality in terms of MSE and PSNR, which are counted using the generated holographic images corresponding to 10,000 target images from the testing dataset of handwritten digits. Considering the inconsistency of the GS algorithm, for each target holographic image, the GS algorithm is run three times, and the smallest MSE value and the largest PSNR value are recorded as the final statistical results.

The mean MSE and PSNR values of our VAE-cGAN evaluated with the whole testing dataset are 0.0382 and 14.33, respectively. Compared with 0.0407 (mean-MSE) and 14.05 (mean-PSNR) of the GS algorithm, our VAE-cGAN shows a better capability in searching the global optimum than the GS algorithm.

### 4. CONCLUSION

We propose a new intelligent non-iterative approach (VAE-cGAN) based on deep learning methods for metasurface holograms. The usage of an unsupervised VAE structure makes our system easily trained from scratch, and the introduction of Wasserstein distance criterion improves the imaging quality of holographic images. After the VAE-cGAN is well trained, we just need to use the generator part to generate the coding patterns (metasurface holograms) corresponding to the target holographic images. The non-iterative structure of the generator enables the realization of holographic imaging with high quality and high efficiency, which are validated by both simulation and experimental results.

When deploying our trained generator on a neural network chip, our system could become a real-time holographical imager to rapidly generate the desired holographic images. It could be expected that our intelligent metasurface hologram system could become an efficient tool for microwave or even optical holograms [69], and more valuable applications could be explored in wireless communications [32], smart EM environment, health monitoring [51,70], and so on.

**Funding.** National Key Research and Development Program of China (2017YFA0700201, 2017YFA0700202, 2017YFA0700203, 2018YFA0701900).

**Disclosures.** The authors declare no conflicts of interest.

### REFERENCES

1. V. G. Veselago, "Electrodynamics of substances with simultaneously negative values of sigma and  $\mu$ ," Sov. Phys. Usp. **10**, 509–514 (1968).
2. R. Zhao, L. Huang, and Y. Wang, "Recent advances in multi-dimensional metasurfaces holographic technologies," PhotoniX **1**, 20 (2020).
3. X. Ding, Z. Wang, G. Hu, J. Liu, K. Zhang, H. Li, B. Ratni, S. N. Burokur, Q. Wu, J. Tan, and C.-W. Qiu, "Metasurface holographic image projection based on mathematical properties of Fourier transform," PhotoniX **1**, 16 (2020).

4. Q. Ma, Q. R. Hong, X. X. Gao, H. B. Jing, C. Liu, G. D. Bai, Q. Cheng, and T. J. Cui, "Smart sensing metasurface with self-defined functions in dual polarizations," *Nanophotonics* **9**, 3271–3278 (2020).
5. X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Broadband light bending with plasmonic nanoantennas," *Science* **335**, 427 (2012).
6. N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *Science* **334**, 333–337 (2011).
7. D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies," *Science* **314**, 977–980 (2006).
8. J. Li and J. B. Pendry, "Hiding under the carpet: a new strategy for cloaking," *Phys. Rev. Lett.* **101**, 203901 (2008).
9. R. Liu, C. Ji, J. J. Mock, J. Y. Chin, T. J. Cui, and D. R. Smith, "Broadband ground-plane cloak," *Science* **323**, 366–369 (2009).
10. Q. Ma, Z. L. Mei, S. K. Zhu, T. Y. Jin, and T. J. Cui, "Experiments on active cloaking and illusion for Laplace equation," *Phys. Rev. Lett.* **111**, 173901 (2013).
11. W. X. Jiang, T. J. Cui, Q. Cheng, J. Y. Chin, X. M. Yang, R. Liu, and D. R. Smith, "Design of arbitrarily shaped concentrators based on conformally optical transformation of nonuniform rational B-spline surfaces," *Appl. Phys. Lett.* **92**, 264101 (2008).
12. Y. Lai, J. Ng, H. Chen, D. Han, J. Xiao, Z.-Q. Zhang, and C. T. Chan, "Illusion optics: the optical transformation of an object into another object," *Phys. Rev. Lett.* **102**, 253902 (2009).
13. L. Chen, Q. Ma, Q. F. Nie, Q. R. Hong, H. Y. Cui, Y. Ruan, and T. J. Cui, "Dual-polarization programmable metasurface modulator for near-field information encoding and transmission," *Photon. Res.* **9**, 116–124 (2021).
14. N. Kundtz and D. R. Smith, "Extreme-angle broadband metamaterial lens," *Nat. Mater.* **9**, 129–132 (2010).
15. W. X. Jiang, C.-W. Qiu, T. C. Han, Q. Cheng, H. F. Ma, S. Zhang, and T. J. Cui, "Broadband all-dielectric magnifying lens for far-field high-resolution imaging," *Adv. Mater.* **25**, 6963–6968 (2013).
16. X. M. Yang, X. Y. Zhou, Q. Cheng, H. F. Ma, and T. J. Cui, "Diffuse reflections by randomly gradient index metamaterials," *Opt. Lett.* **35**, 808–810 (2010).
17. T. J. Cui, M. Q. Qi, X. Wan, J. Zhao, and Q. Cheng, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light Sci. Appl.* **3**, e218 (2014).
18. J. Li, Y. Zhang, J. Li, X. Yan, L. Liang, Z. Zhang, J. Huang, J. Li, Y. Yang, and J. Yao, "Amplitude modulation of anomalously reflected terahertz beams using all-optical active Pancharatnam–Berry coding metasurfaces," *Nanoscale* **11**, 5746–5753 (2019).
19. R. Y. Wu, L. Zhang, L. Bao, L. W. Wu, Q. Ma, G. D. Bai, H. T. Wu, and T. J. Cui, "Digital metasurface with phase code and reflection-transmission amplitude code for flexible full-space electromagnetic manipulations," *Adv. Opt. Mater.* **7**, 1801429 (2019).
20. Q. Ma, Q. R. Hong, G. D. Bai, H. B. Jing, R. Y. Wu, L. Bao, Q. Cheng, and T. J. Cui, "Editing arbitrarily linear polarizations using programmable metasurface," *Phys. Rev. Appl.* **13**, 021003 (2020).
21. Q. Ma, C. B. Shi, G. D. Bai, T. Y. Chen, A. Noor, and T. J. Cui, "Beam-editing coding metasurfaces based on polarization bit and orbital-angular-momentum-mode bit," *Adv. Opt. Mater.* **5**, 1700548 (2017).
22. Q. Ma, L. Chen, H. B. Jing, Q. R. Hong, H. Y. Cui, Y. Liu, L. Li, and T. J. Cui, "Controllable and programmable nonreciprocity based on detachable digital coding metasurface," *Adv. Opt. Mater.* **7**, 1901285 (2019).
23. G. Ding, K. Chen, X. Luo, J. Zhao, T. Jiang, and Y. Feng, "Dual-helicity decoupled coding metasurface for independent spin-to-orbital angular momentum conversion," *Phys. Rev. Appl.* **11**, 044043 (2019).
24. J. Han, L. Li, H. Yi, and Y. Shi, "1-bit digital orbital angular momentum vortex beam generator based on a coding reflective metasurface," *Opt. Mater. Express* **8**, 3470–3478 (2018).
25. Q. Zheng, Y. Li, Y. Han, M. Feng, Y. Pang, J. Wang, H. Ma, S. Qu, and J. Zhang, "Efficient orbital angular momentum vortex beam generation by generalized coding metasurface," *Appl. Phys. A* **125**, 136 (2019).
26. T. J. Cui, L. Li, S. Liu, Q. Ma, and Q. Cheng, "Information metamaterial systems," *iScience* **23**, 101403 (2020).
27. L. Li and T. J. Cui, "Information metamaterials - from effective media to real-time information processing systems," *Nanophotonics* **8**, 703–724 (2019).
28. T. J. Cui, S. Liu, and L. Zhang, "Information metamaterials and metasurfaces," *J. Mater. Chem. C* **5**, 3644–3668 (2017).
29. Q. Ma and T. J. Cui, "Information metamaterials: bridging the physical world and digital world," *PhotonIX* **1**, 1 (2020).
30. L. Zhang, X. Q. Chen, S. Liu, Q. Zhang, J. Zhao, J. Y. Dai, G. D. Bai, X. Wan, Q. Cheng, G. Castaldi, V. Galdi, and T. J. Cui, "Space-time-coding digital metasurfaces," *Nat. Commun.* **9**, 4334 (2018).
31. T. J. Cui, S. Liu, G. D. Bai, and Q. Ma, "Direct transmission of digital message via programmable coding metasurface," *Research* **2019**, 2584509 (2019).
32. H. Zhao, Y. Shuang, M. Wei, T. J. Cui, P. D. Hougne, and L. Li, "Metasurface-assisted massive backscatter wireless communication with commodity Wi-Fi signals," *Nat. Commun.* **11**, 3926 (2020).
33. Y. Shuang, H. Zhao, W. Ji, T. J. Cui, and L. Li, "Programmable high-order OAM-carrying beams for direct-modulation wireless communications," *IEEE J. Emerg. Sel. Top. Circuits Syst.* **10**, 29–37 (2020).
34. G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath, and B. Kingsbury, "Deep neural networks for acoustic modeling in speech recognition: the shared views of four research groups," *IEEE Signal Process. Mag.* **29**, 82–97 (2012).
35. G. Pironkov, S. U. N. Wood, and S. Dupont, "Hybrid-task learning for robust automatic speech recognition," *Comput. Speech Lang.* **64**, 101103 (2020).
36. A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing* (2013), pp. 6645–6649.
37. K. Lu, "Intelligent recognition system for high precision image significant features in large data background," in *Cyber Security Intelligence and Analytics* (2020), pp. 1056–1062.
38. T. Tong, X. Mu, L. Zhang, Z. Yi, and P. Hu, "MBVCNN: joint convolutional neural networks method for image recognition," *AIP Conf. Proc.* **1839**, 020091 (2017).
39. H. Sun, Q. Zhang, H. Wang, and A. Li, "Research of images recognition method based on RBF neural network," in *7th International Conference on System of Systems Engineering* (2012), pp. 24–26.
40. K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. J. C. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv:1406.1078v3* (2014).
41. C. Escalano, M. R. Costa-Jussà, and J. A. R. Fonollosa, "From bilingual to multilingual neural-based machine translation by incremental training," in *57th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (2019), pp. 236–242.
42. S. Kwon, B. H. Go, and J. H. Lee, "A text-based visual context modulation neural model for multimodal machine translation," *Pattern Recognit. Lett.* **136**, 212–218 (2020).
43. C. He, Y. Wan, Y. Gu, and F. L. Lewis, "Integral reinforcement learning-based multi-robot minimum time-energy path planning subject to collision avoidance and unknown environmental disturbances," *IEEE Control Syst. Lett.* **5**, 983–988 (2021).
44. B. Sangiovanni, G. P. Incremona, M. Piastra, and A. Ferrara, "Self-configuring robot path planning with obstacle avoidance via deep reinforcement learning," *IEEE Control Syst. Lett.* **5**, 397–402 (2021).
45. J. Yoo, D. Jang, H. J. Kim, and K. H. Johansson, "Hybrid reinforcement learning control for a micro quadrotor flight," *IEEE Control Syst. Lett.* **5**, 505–510 (2021).
46. Q. Zhang, X. Wan, S. Liu, J. Y. Yin, L. Zhang, and T. J. Cui, "Shaping electromagnetic waves using software-automatically-designed metasurfaces," *Sci. Rep.* **7**, 3588 (2017).
47. L. L. Li, H. X. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alu, C. W. Qiu, and T. J. Cui, "Machine-learning reprogrammable metasurface imager," *Nat. Commun.* **10**, 1082 (2019).
48. T. Qiu, X. Shi, J. Wang, Y. Li, S. Qu, Q. Cheng, T. Cui, and S. Sui, "Deep learning: a rapid and efficient route to automatic metasurface design," *Adv. Sci.* **6**, 1900128 (2019).

49. Q. Ma, G. D. Bai, H. B. Jing, C. Yang, L. Li, and T. J. Cui, "Smart metasurface with self-adaptively reprogrammable functions," *Light Sci. Appl.* **8**, 98 (2019).
50. H. Li, H. Zhao, M. Wei, H. Ruan, Y. Shuang, T. J. Cui, P. del Hougne, and L. Li, "Intelligent electromagnetic sensing with learnable data acquisition and processing," *Patterns* **1**, 100006 (2020).
51. L. Li, Y. Shuang, Q. Ma, H. Li, H. Zhao, M. Wei, C. Liu, C. Hao, C.-W. Qiu, and T. J. Cui, "Intelligent metasurface imager and recognizer," *Light Sci. Appl.* **8**, 97 (2019).
52. L. Li, T. J. Cui, W. Ji, S. Liu, J. Ding, X. Wan, Y. B. Li, M. Jiang, C.-W. Qiu, and S. Zhang, "Electromagnetic reprogrammable coding-metasurface holograms," *Nat. Commun.* **8**, 197 (2017).
53. J. Wu, Z. Wang, L. Zhang, Q. Cheng, S. Liu, S. Zhang, J. Song, and C. T. Jun, "Anisotropic metasurface holography in 3D space with high resolution and efficiency," *IEEE Trans. Antennas Propag.* **69**, 302–316 (2020).
54. C. Liu, Q. Ma, L. Li, and T. J. Cui, "Work in progress: intelligent metasurface holograms," in *1st ACM International Workshop on Nanoscale Computing, Communication, and Applications* (2020), pp. 45–48.
55. Z. Y. Zhou, J. Xia, J. Wu, C. L. Chang, X. Ye, S. G. Li, B. T. Du, H. Zhang, and G. D. Tong, "Learning-based phase imaging using a low-bit-depth pattern," *Photon. Res.* **8**, 1624–1633 (2020).
56. R. Lopez, J. Regier, M. I. Jordan, and N. Yosef, "Information constraints on auto-encoding variational Bayes," in *32nd Conference on Neural Information Processing Systems (NeurIPS)* (2018), pp. 6114–6125.
57. I. Deshpande, Z. Zhang, and A. Schwing, "Generative modeling using the sliced Wasserstein distance," in *IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 3483–3491.
58. M. Arjovsky and L. Bottou, "Towards principled methods for training generative adversarial networks," in *5th International Conference on Learning Representations (ICLR)* (2017), pp. 1–17.
59. V. M. Panaretos and Y. Zemely, "Statistical aspects of Wasserstein distances," arXiv:1806.05500v3 (2019).
60. M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *34th International Conference on Machine Learning (ICML)* (2017), pp. 214–223.
61. I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of wasserstein GANs," in *31st Annual Conference on Neural Information Processing Systems (NIPS)* (2017), pp. 5768–5778.
62. M. Mirza, "Conditional generative adversarial nets," arXiv: 1411.1784v1 (2014).
63. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 5967–5976.
64. T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs," in *31st Meeting of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2018), pp. 8798–8807.
65. K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," in *21st ACM Conference on Computer and Communications Security (CCS)* (2016), pp. 630–645.
66. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *32nd International Conference on Machine Learning (ICML)* (2015), pp. 448–456.
67. D. Ulyanov and A. Vedaldi, "Instance normalization: the missing ingredient for fast stylization," arXiv:1607.08022v3 (2017).
68. Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST database of handwritten digits," <http://yann.lecun.com/exdb/mnist/> (2012).
69. X. Zou, G. Zheng, Q. Yuan, W. Zang, R. Chen, T. Li, L. Li, S. Wang, Z. Wang, and S. Zhu, "Imaging based on metalenses," *PhotonX* **1**, 2 (2020).
70. A. Ahad, M. Tahir, and K. A. Yau, "5G-based smart healthcare network: architecture, taxonomy, challenges and future research directions," *IEEE Access* **7**, 100747 (2019).

# PHOTONICS Research

## Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images

ZAFRAN HUSSAIN SHAH,<sup>1</sup> MARCEL MÜLLER,<sup>2</sup> TUNG-CHENG WANG,<sup>2</sup> PHILIP MAURICE SCHEIDIG,<sup>1</sup> AXEL SCHNEIDER,<sup>1</sup> MARK SCHÜTTPELZ,<sup>2</sup> THOMAS HUSER,<sup>2,3</sup> AND WOLFRAM SCHENCK<sup>1,4</sup>

<sup>1</sup>Faculty of Engineering and Mathematics, Bielefeld University of Applied Sciences, 33619 Bielefeld, Germany

<sup>2</sup>Faculty of Physics, Bielefeld University, 33615 Bielefeld, Germany

<sup>3</sup>e-mail: thomas.huser@physik.uni-bielefeld.de

<sup>4</sup>e-mail: wolfram.schenck@fh-bielefeld.de

Received 1 December 2020; revised 14 January 2021; accepted 26 January 2021; posted 29 January 2021 (Doc. ID 416437); published 14 April 2021

**Super-resolution structured illumination microscopy (SR-SIM) provides an up to twofold enhanced spatial resolution of fluorescently labeled samples.** The reconstruction of high-quality SR-SIM images critically depends on patterned illumination with high modulation contrast. Noisy raw image data (e.g., as a result of low excitation power or low exposure time), result in reconstruction artifacts. Here, we demonstrate deep-learning based SR-SIM image denoising that results in high-quality reconstructed images. A residual encoding-decoding convolutional neural network (RED-Net) was used to successfully denoise computationally reconstructed noisy SR-SIM images. We also demonstrate the end-to-end deep-learning based denoising and reconstruction of raw SIM images into high-resolution SR-SIM images. Both image reconstruction methods prove to be very robust against image reconstruction artifacts and generalize very well across various noise levels. The combination of computational image reconstruction and subsequent denoising via RED-Net shows very robust performance during inference after training even if the microscope settings change. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416437>

### 1. INTRODUCTION

Fluorescence microscopy remains one of the most powerful tools for imaging cell biology samples because of its ability to specifically label molecular structures and to visualize them in different fluorescence color channels. It also offers exceptionally high sensitivity and can visualize molecular processes even below the optical diffraction limit. Several methods that enable imaging of fluorescently labeled samples down to the nanoscale have been developed during the last two decades [1]. Super-resolution structured illumination microscopy (SR-SIM) is a particularly compelling method, because it works with the majority of samples and fluorophores commonly used in cell biology without imposing specific requirements on sample preparation [2,3]. Therefore, it can even be applied to living samples [4–12]. SR-SIM in its most common form uses a series of sinusoidal illumination patterns with a pattern periodicity at or near the diffraction limit. This patterned excitation light is phase-shifted laterally and rotated to different discrete angles to acquire a series of raw images, which are then passed on to an image reconstruction algorithm to obtain the final super-resolved SR-SIM image [13]. Several implementations of

reconstruction algorithms that operate in the frequency space were developed and a number of open access tools are now available that aim to enhance the speed and spatial resolution, and minimize reconstruction artifacts in the final reconstructed image (i.e., fairSIM [14], OpenSIM [15], SIMToolbox [16], and CC-SIM [17]). The most common property among all these image reconstruction algorithms is that they require a series of high-quality raw images to be able to reconstruct a high-quality super-resolution image. However, reconstruction in the frequency space fails to reliably reconstruct SR-SIM images if the signal-to-noise ratio (SNR) is too low (e.g., because the laser excitation power level was too low), the sample exposure time chosen was too short, or the sample has already undergone irreversible photobleaching [18]. Recently, several developments that aim to reduce reconstruction artifacts from SR-SIM images have been undertaken. Huang *et al.* used a fully analytical approach to reduce noise and minimize reconstruction artifacts using the Hessian matrix theory [10]. Hoffman and Betzig proposed the reconstruction of SIM images in lower pixel count tiles with subsequent merger to reduce reconstruction artifacts [19]. Jin *et al.*, on the other hand, used deep neural networks to reconstruct the cropped

regions of SR-SIM images [20]. Christensen *et al.* used a deep learning architecture for the reconstruction from synthetic raw SIM images [21] with subsequent testing on real microscope images. Ling *et al.* relied on a special type of convolutional neural network, a CycleGAN, for the same purpose [22]. And Weigert *et al.* used deep learning algorithms to enhance isotropic resolution and the signal-to-noise ratio (SNR) of fluorescence microscopy images in general [23].

Here, we provide a comprehensive study on several deep-learning based approaches to denoise and reconstruct SR-SIM images. First, as starting point, we include an end-to-end deep learning architecture and workflow that is related to the existing literature combining SIM and deep learning [20–22] in that a single network is used that receives raw SIM images as input and produces a super-resolution image as the output. This approach is named super-resolution residual encoder-decoder structured illumination microscopy (SR-REDSIM) in the following discussion. In this case, the entire SIM reconstruction process is performed by the deep convolutional neural network.

Second, for what we believe, to the best of our knowledge, is an entirely novel workflow, we combine classical computational SIM reconstruction with a deep learning network. This workflow is called a residual encoder-decoder fairSIM (RED-fairSIM). RED-fairSIM is a combination of the fairSIM image reconstruction package [14], which performs image reconstruction using commonly used frequency domain algorithms and a deep convolutional neural network for subsequent artifact reduction and denoising. Finally, we also show the results for a workflow where deep learning is first applied to the raw SIM images for denoising with subsequent classical computational SIM reconstruction; thus, exactly the other way around, as in RED-fairSIM. This approach is named preRED-fairSIM.

For the main network architecture for all three approaches we use the residual encoder-decoder network (RED-Net) [24], which is comparatively lightweight. For training, raw image data directly from the microscope is used, which offers a straightforward, practical solution without the need to create synthetic data or carry out substantive preprocessing in contrast to the existing literature. We found that the first two methods are robust in their ability to significantly improve the quality of the reconstructed SR-SIM images. We show, for what we believe is the first time, that the trained networks of these two methods generalize well to real microscopic SIM data with different SNRs. The novel RED-fairSIM workflow shows superior performance in this regard. Furthermore, we also demonstrate that RED-fairSIM generalizes better even if the microscope settings are changed after training.

## 2. MATERIALS AND METHODS

### A. Training Data Generation

Deep learning methods rely on training data, which, in our case, consist of noisy raw SIM images for the input and ideally noise and artifact free, super-resolved SIM reconstructions for the output. Thus, we first need to generate such a data set that is large enough to effectively train the network, but also captures all aspects of the SIM imaging process (sample behavior,

instrument imperfections, and data processing artifacts) well enough.

In principle, the data acquisition process of a SIM microscope can be simulated. In this case, the expected output represents the ground truth data upon which the simulation is based, and which is known without SNR or resolution limits. In addition, the amount of available training data would only be limited by the processing time, as the generation would be fully automated and not rely on access to a microscope system. However, we decided against this pure *in silico* approach. Although the basic effects of structured illumination, Poisson-distributed noise, and even basic optical imperfections are rather easy to simulate, modeling the response of a full structured illumination microscope correctly is very complex. Additionally, such a simulation would likely have to be adjusted to reflect the properties of a specific SR-SIM instrument to capture changes (e.g., when switching to a different manufacturer or even a specific installation of an SR-SIM microscope). The same argument holds true for the fluorescent samples themselves. Although some simulations that provide perfect ground-truth data exist (e.g., for single molecule localization microscopy [25,26]), they, again, do not capture all of the variability found in real-world samples.

The option chosen for data generation for the work presented here is to use real microscope data from standard biological samples. This approach naturally captures all aspects and imperfections of the samples and of the specific instrument in question [23], but also poses constraints. Because data collection requires both instrument time and manual sample preparation and handling, the amount of training data is naturally limited. There also is no perfect ground truth available. To acquire the high-quality reference images presented to the networks as desired output, we adjusted the instrument to provide high-SNR raw frames and processed those with the classical, frequency-domain-based image reconstruction algorithm. While these images are low in noise and reconstruction artifacts, they are never completely devoid of them. For this reason, we refer to them by the term “reference image” instead of “ground truth.” To acquire noisy, low-SNR images as input, the samples were then photo-bleached by continued exposure and acquisition of raw SIM images, which naturally reduces the fluorescent light output over time and results in a series of images with steadily decreasing SNR. (See Section 2.C for details.)

### B. Sample Preparation and Data Acquisition

U2OS cells were cultured in DMEM supplemented with 10% FBS and grown on round coverslips of  $170 \pm 5$  nm thickness (No. 1.5H). Cells were fixed with 4% PFA for 15 min, followed by PBS washes, and permeabilization with 0.5% Triton-X100 for 3 min. Another two rounds of PBS washes were done prior to blocking with 3% BSA. For the immunolabeling, the microtubuli, cells were stained with anti-tubulin Ab (Invitrogen Cat. No. 322500) 1:400 for 2 h at room temperature, followed by a PBS wash and one additional hour of incubation with Alexa 488-conjugated anti-mouse IgG 1:400. Cells were then briefly washed with PBS before Vectashield was applied to embed the coverslip onto the standard slide glass for imaging.

The DeltaVision|OMX (GE Healthcare, Chicago, IL, USA) was used to acquire 3D-SIM raw images. A total of 101 randomly selected fields of view were acquired by exposing each field-of-view to the full laser power of the 488 nm excitation laser and the exposure time was set at 20 ms for each of the 15 raw image frames. A total of 200 image repetitions were collected at each position without delays. Taking into account the camera readout and pattern switching time, acquiring 15 raw SIM images, making up one timestamp in the raw image stack, takes approximately 375 ms.

### C. Data Preprocessing

In our work, we used 101 different cell structures (fields of view). Out of these, 81 were selected for the training data and the remaining 20 for the test data. Each cell structure was captured for 200 repetitions. During each repetition, 15 frames were recorded, iterating the phase and orientation of the sinusoidal SIM illumination pattern, yielding an image stack of size  $15 \times 512 \times 512$  (frames, width, height). During this time-lapse acquisition, the samples underwent photo-bleaching, which reduces the amount of active fluorescent emitters and thus the amount of emitted photons. Therefore, less and less light is captured and the SNR steadily decreases during the acquisition of such a time series.

The cell structures from the timestamps 175 to 200 are therefore considered as noisy training and test input while the samples from timestamp 0 are considered as clean output images. All 15 clean raw SIM images of the 101 cell structures from timestamp 0 are used to reconstruct high-resolution reference SIM images of size  $1024 \times 1024$  pixels by using fairSIM, which employs a classic frequency-domain-based reconstruction. In this work, the input dimension is  $15 \times 512 \times 512$  (frames, width, height) whereas the output dimension is  $1024 \times 1024$  (width, height) pixels. A total of 2525 samples were further divided into training and test data. The training data contains 2025 images of the first 81 cell structures; the test data is composed of 500 test images that are created from the remaining 20 cell structures.

The only preprocessing step involved in our work is the linear scaling of the training and test data to match the overall brightness between the input and output. In addition, we tested an image augmentation approach to double the amount of training data by rotating each image by an angle of  $180^\circ$ .

The data from each time series over 200 repetitions were subdivided into different noise levels. Noise level 0 stands for the highest SNR in our data at timestamp 0. In our work, this is our reference data. The image data from timestamps 175–200 represent the highest noise level 4, the data from timestamps 125–150 represent noise level 3, the data from timestamps 75–100 are noise level 2, and the data from timestamps 25–50 are noise level 1. In this study, data from noise level 4 are only used in the training process whereas data from noise levels 1, 2, 3, and 4 are used in the test phase.

### D. Architecture and Training of SR-REDSIM

In the first deep-learning based SR-SIM image reconstruction method, named SR-REDSIM, the reconstruction and denoising of noisy raw SIM images are both performed by a single deep learning model. This model is a modified version of the

RED-Net. RED-Net is an encoding-decoding framework with symmetric convolutional-deconvolutional layers along with skip-layer connections. It was previously used to accomplish different image restoration tasks such as image denoising, image super-resolution, and image inpainting [24]. The original RED-Net architecture is only composed of encoding-decoding blocks with the size of the network input being the same as the size of the network output. Therefore, super-resolution with this architecture has to rely on explicit image pre-upsampling [24]. In contrast, our modified RED-Net architecture contains an additional upsampling block after the encoding-decoding blocks. This upsampling block inside our model has the advantage that the input images are first denoised in their lower-dimensional space, which reduces the training time and effort.

Most of the super-resolution architectures such as the enhanced deep super-resolution network (EDSR) [27] or the residual channel attention network (RCAN) [28] are very deep and require a significant amount of training data. In comparison, our architecture is comparably lightweight.

SR-REDSIM is based on a modified version of RED-Net. The complete pipeline of this approach is shown in Fig. 1(a), whereas the architecture of SR-REDSIM and details about the model parameters are given in Fig. 2(a). The SR-REDSIM architecture consists of three blocks: the encoder, the decoder, and the upsampling block. SR-REDSIM contains a total of 44 convolutional and deconvolutional layers with symmetric skip connections. The encoder block is composed of 21 convolutional layers, whereas the decoder contains 21 deconvolutional layers. The upsampling block consists of two deconvolutional layers that perform the upsampling task by adjusting the size of the stride. The SR-REDSIM model provides the best results after training the model for 100 epochs. The SR-REDSIM model is trained only with high-level noise data from timestamps 175 to 200. During the training process, the ADAM optimizer and the L2 loss function, also known as least squares error, are used, so

$$\text{Loss} = \sum_{i=1}^n (y_i - \hat{y}_i)^2. \quad (1)$$

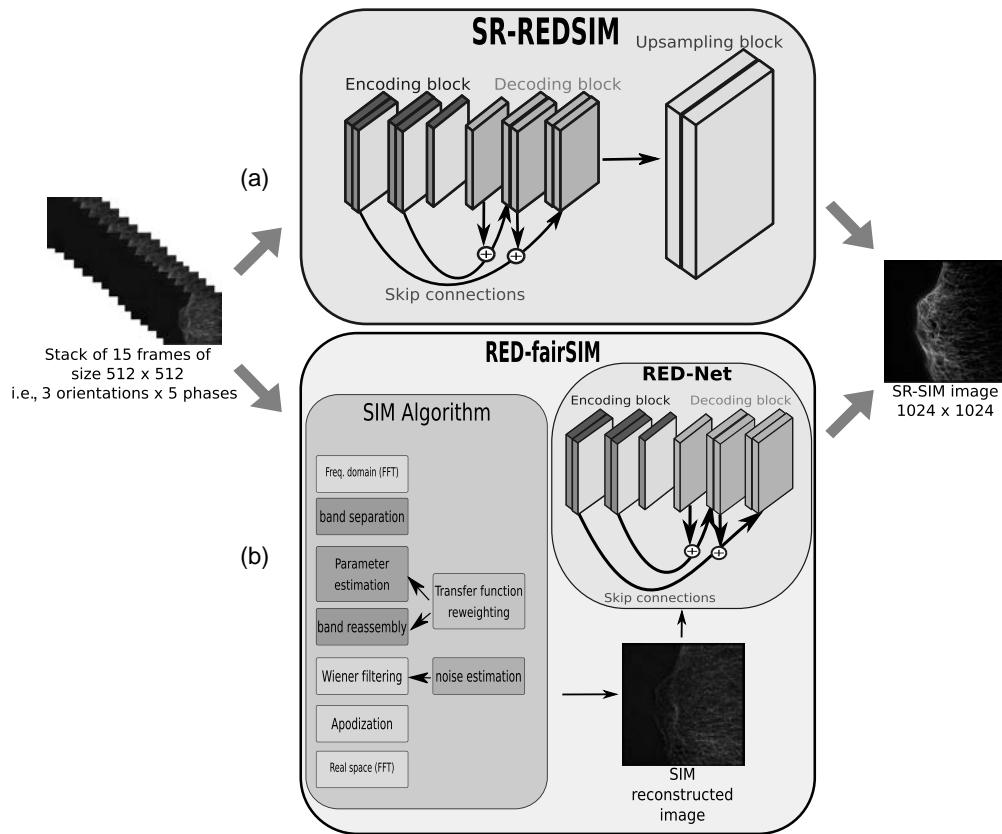
In Eq. (1),  $y_i$  represents the true pixel intensity,  $\hat{y}_i$  represents the predicted pixel intensity, and  $n$  is the number of pixels (in the image).

### E. Architecture and Training of RED-fairSIM

The RED-fairSIM is combination of fairSIM and RED-Net. The pipeline of RED-fairSIM is shown in Fig. 1(b). In this approach, first fairSIM is used to transform the raw SIM images into a super-resolved output image by employing a classic, frequency-domain SIM reconstruction algorithm. This output image now contains noise, which, due to the frequency-domain algorithm, takes a SIM-specific form, and might show other reconstruction artifacts. It is subsequently processed by RED-Net.

#### 1. fairSIM

The fairSIM reconstruction is performed in three steps: parameter estimation, reconstruction, and filtering. Mathematical and algorithmic details are provided in the original publication



**Fig. 1.** Schematics of the deep learning CNN architecture of the two different SR-SIM image denoising (RED-fairSIM) and image denoising and reconstruction (SR-REDSIM) methods. In both approaches, a stack of 15 raw (noisy) SIM images (three angles with five phases each) is used as input. The output is the reconstructed SR-SIM image. (a) SR-REDSIM is composed of three main blocks. The encoding block contains mainly the convolutional layers whereas the decoding block consists of deconvolutional layers and the upsampling block of deconvolutional upsampling layers. (b) In the RED-fairSIM method, fairSIM is first used to computationally reconstruct noisy SR-SIM images that are then further propagated into the RED-Net for denoising. The architecture of RED-Net is composed of the encoder and the decoder blocks.

[14]. A synthetic optical transfer function, with  $\text{NA} = 1.4$ ,  $\lambda = 525 \text{ nm}$ , and  $\alpha = 0.31$  ( $\alpha$  is a compensation parameter [14,29]), is used. A total of 500 counts of background are subtracted per pixel. SIM reconstruction parameters (pattern orientation and global phase, for example) are automatically determined by the fairSIM standard in an iterative cross-correlation approach. Filter parameters are set to a generalized Wiener filter strength of  $w = 0.05$ , apodization is set at  $1.9 \times$  the resolution limit with a bend of 0.8, and a notch-style filter is implemented as OTF attenuation with a strength of 0.995 and an FWHM of  $1.2 \mu\text{m}^{-1}$  is in use. For the detailed meaning and influence of these parameters, please refer to the fairSIM source code, its accompanying publication [14], and this general guide to SIM reconstruction parameters [29].

## 2. RED-Net Architecture

The architecture of RED-Net used in RED-fairSIM consists of 15 convolutional and 15 deconvolutional layers along with symmetric skip connections, as shown in Fig. 2(b). The output of fairSIM is propagated into the RED-Net to denoise the reconstructed noisy sample. During the training phase, the noisy SR-SIM images of size  $1024 \times 1024$  pixels along with the reference SR-SIM images of the same size are used as input-output

pairs for the RED-Net. The network is trained for 100 epochs with the ADAM optimizer and L2 loss.

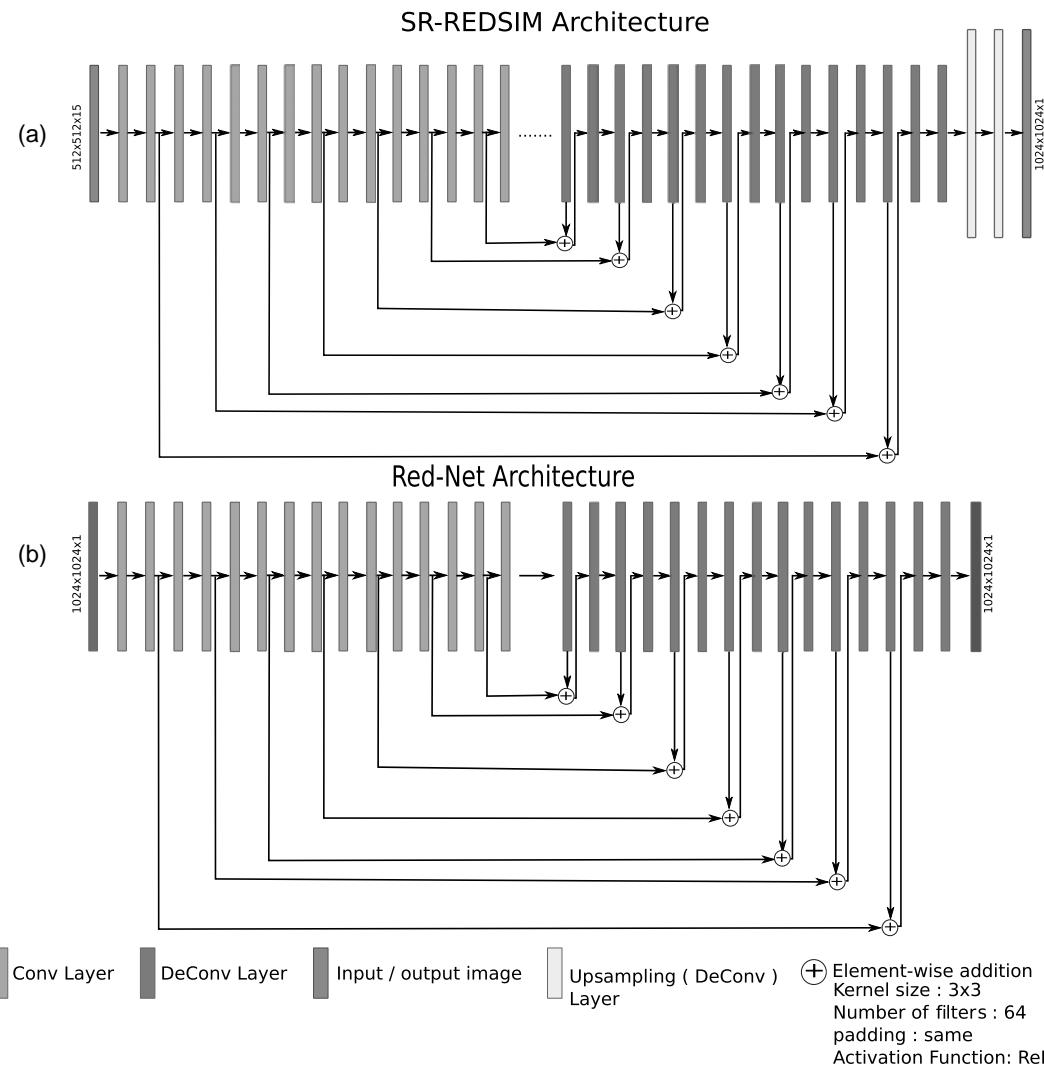
## F. Architecture and Training of U-Net-fairSIM

U-Net is also a popular deep learning architecture that is extensively used in the domain of image restoration such as image denoising and super-resolution [23,30,31]. In U-Net-fairSIM, we simply replaced the RED-Net from the RED-fairSIM approach with the U-Net architecture. The U-Net is also trained for 100 epochs with the ADAM optimizer and L2 loss.

## 3. RESULTS

### A. SR-REDSIM: SR-SIM Image Denoising and Reconstruction Using the Super-resolution REDSIM Method

SR-REDSIM is an entirely deep-learning based, end-to-end method. The complete pipeline of SR-REDSIM is shown in Fig. 1(a) and the architecture of SR-REDSIM, as shown in Fig. 2(a), is explained in more detail in Section 2.D. During the training process, we used all 15 raw noisy SIM images (three angles with five phases each) of size  $512 \times 512$  pixels [i.e., stack dimensions were  $15 \times 512 \times 512$  (frames, width, height)] as input along with the reconstructed super-resolved SIM image



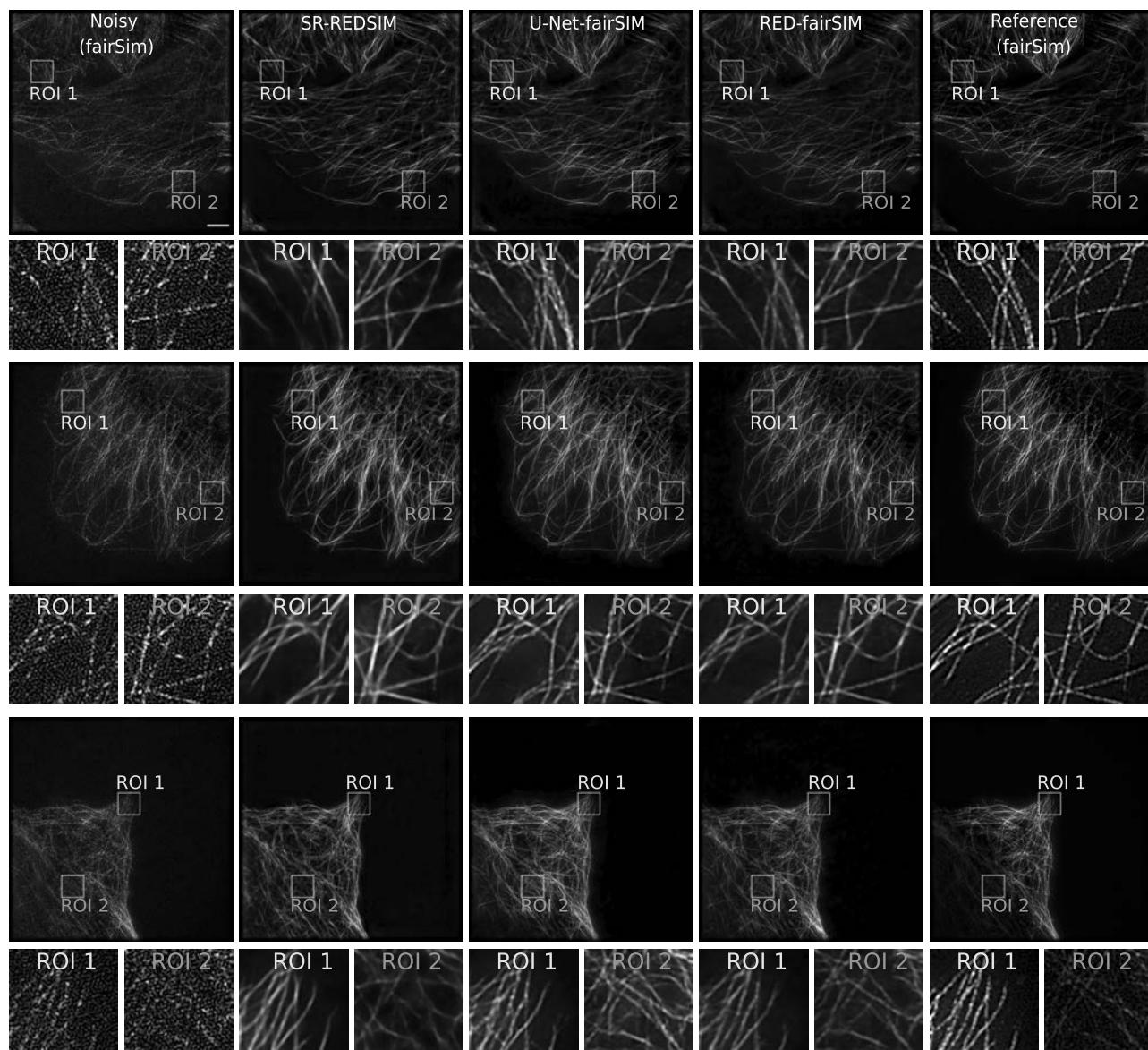
**Fig. 2.** The architecture of the networks used in this work. (a) SR-REDSIM architecture is composed of three different blocks. The encoding and decoding blocks contain 21 convolutional and deconvolutional layers, respectively, whereas the upsampling blocks consist only of two upsampling layers. This architecture was used in the SR-REDSIM method to denoise and reconstruct the raw SIM images. (b) The complete RED-Net architecture contains 15 convolutional and 15 deconvolutional layers along with the additive symmetric skip connection layers. This architecture was used in the RED-fairSIM and preRED-fairSIM methods for denoising (in preRED-fairSIM, the input and output have a size of only  $512 \times 512 \times 1$ ).

of size  $1024 \times 1024$  pixels as output. The output was generated by the fairSIM software from raw SIM images recorded with the highest SNRs, while the input images were taken from noise level 4. Note that Section 2.C offers an explanation of the noise levels. The trained network was tested afterward on unseen test data from noise level 4. The super-resolution images obtained during this test are depicted in column 2 of Fig. 3, whereas columns 1 and 5 show the results of noisy fairSIM (reconstructed by fairSIM from noisy raw SIM images; noise level 4) and reference fairSIM (reconstructed by fairSIM from raw SIM images with the highest SNR). The comparison of these images and of specific regions of interest (ROIs) between fairSIM in Fig. 3 (column 1, all rows) and SR-REDSIM (column 2, all rows) clearly shows that the noise is completely removed by SR-REDSIM. However, in the reconstruction by SR-REDSIM, fine cell structures are partly

suppressed compared to the reference output (compare column 2, row 2, ROI 1 in Fig. 3 with column 5, row 2, ROI 1). In rows 3/4 and 5/6 of Fig. 3, the structure of the cell is well denoised and reconstructed by SR-REDSIM. Moreover, the evaluation of the SR-REDSIM method on the basis of peak SNR (PSNR) and structural similarity index measurement (SSIM) [32] values in Table 1 shows a significant improvement compared to fairSIM.

#### B. RED-fairSIM: SR-SIM Reconstruction of Noisy Input Data by Using a Combination of fairSIM and RED-Net

One of the well-known open-source reconstruction algorithm implementations is fairSIM. It is widely used for super-resolution tasks among the other tools in SIM microscopy. However, it cannot reconstruct a clean, high-quality super-resolution



**Fig. 3.** Super-resolution SIM (SR-SIM) images of three different test samples (U2OS osteosarcoma cells, tubulin cytoskeleton labeled with anti-tubulin-Alexa488) at high-level noise (time stamps 175–199; noise level 4). Each column represents a different reconstruction approach: fairSIM (first column), SR-REDSIM (second column), U-Net-fairSIM (third column), and RED-fairSIM (fourth column). The fifth column depicts the reconstructed reference images which were generated by applying fairSIM image reconstruction to high SNR image data at noise level 0 (lowest noise level; i.e., timestamp 0). All reconstructed SR-SIM images have  $1024 \times 1024$  pixels. The first, third, and fifth rows correspond to the full-size SR-SIM images, whereas the second, fourth, and sixth rows depict magnified ROIs of the white squares (bounding boxes) indicated in the full-size images. The extracted ROIs of size  $100 \times 100$  pixels were upsampled to  $300 \times 300$  pixels using bicubic interpolation for illustration purposes. Scale bar:  $4 \mu\text{m}$ .

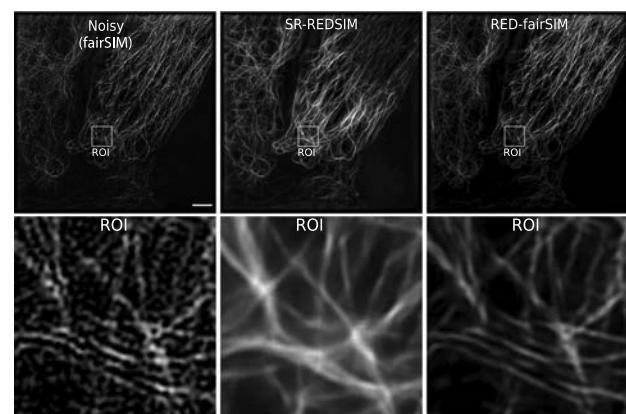
**Table 1.** Mean PSNR and SSIM Values along with the Standard Deviation for All Methods Calculated on 500 Test Images with Respect to the Reference (Highest SNR) Images

	Mean PSNR (STD) and SSIM (STD) Values at Different Noise Levels							
	Noise level 1		Noise level 2		Noise level 3		Noise level 4	
	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)
fairSIM	28.84(2.47)	0.69(0.10)	25.83(1.70)	0.45(0.09)	24.12(1.64)	0.32(0.07)	23.61(1.54)	0.29(0.07)
SR-REDSIM	26.73(1.93)	0.69(0.07)	26.65(1.55)	0.64(0.07)	26.66(1.73)	0.64(0.07)	26.62(1.79)	0.69(0.09)
U-Net-fairSIM	27.44(1.71)	0.75(0.08)	27.23(1.76)	0.69(0.09)	26.85(1.70)	0.65(0.09)	26.80(1.65)	0.68(0.10)
RED-fairSIM	28.75(1.86)	0.80(0.07)	28.67(1.99)	0.75(0.08)	28.18(2.06)	0.70(0.09)	27.97(2.01)	0.71(0.09)

image from noisy raw SIM images. In the RED-fairSIM method, fairSIM is first used to reconstruct the SIM samples of noisy raw SIM images and then RED-Net is used to denoise the output of fairSIM and to generate high-quality super-resolution images. During the reconstruction process, the stack of 15 noisy images (three angles, five phases) of size  $512 \times 512$  pixels is again propagated into the fairSIM reconstruction algorithm that then generates a single noisy reconstructed image of size  $1024 \times 1024$  pixels. The noise and artifacts found in these images do not follow a typical distribution (e.g., Poisson or Gaussian), but have a distinct form that comes from the frequency-based reconstruction algorithm. This single noisy reconstructed image is further passed into the RED-Net architecture to achieve the final result, which can be seen in Fig. 3 (column 4). The complete pipeline of the RED-fairSIM method can be seen in Fig. 1(b) and the architecture of RED-Net is shown in Fig. 2(b). The parameters that were used to generate the SIM reconstructed samples from the raw SIM images are explained in Section 2.E.1. RED-Net was trained in a supervised way where the input–output pairs contain the noisy and reference reconstructed images.

The performance of this method on the unseen test samples is the best among our experiments with respect to PSNR and SSIM values, as shown in Table 1, as well as visually. The ROIs in Fig. 3 show clearly that the output images generated by RED-fairSIM are of high quality with fine details and smooth lines. They are superior compared to the noisy fairSIM, the reference fairSIM, and the SR-REDSIM. Even the artifacts introduced by fairSIM in the reference images are completely removed by RED-fairSIM. It might appear as if the contrast in the denoised SR-REDSIM and RED-fairSIM images in Fig. 3 is weaker when compared to the reference image. This, however, is not the case. The images shown in Fig. 3 are original image data as produced by the various denoising or reconstruction methods without further image processing (i.e. contrast adjustments). Part of the apparent higher contrast in the reference image can also be attributed to graininess of the image, which is caused by uneven antibody staining of the tubulin filaments (a typical phenomenon of antibody staining). It should be noted that the DL-based denoising methods also (at least partly) remove this unevenness.

Furthermore, preliminary tests of RED-fairSIM and SR-REDSIM concerning their ability to generalize to different SIM imaging conditions were carried out, as shown in Fig. 4. As before, U2OS cells were stained for microtubuli, but a dark-red dye with illumination shifted to 642 nm was used, which subsequently also shifts the spatial frequencies of the illumination pattern. The RED-fairSIM approach is able to denoise these images and remove SIM reconstruction artifacts, while the SR-REDSIM approach creates heavy ghosting artifacts. This is unsurprising, as in the case of SR-REDSIM, all specific properties of the SIM pattern (spatial frequencies, orientation, phases) are learned by the network. In the RED-fairSIM approach, parameters specific to the SIM pattern are absorbed by the classic, frequency-domain-based reconstruction, and only reconstruction artifacts are carried into the network. Those artifacts might still depend on the SIM imaging parameters, so



**Fig. 4.** These SR-SIM images show the difference between the output of the SR-REDSIM and RED-fairSIM methods when applied to imaging conditions that the underlying network was not trained for. To evaluate the generalization capabilities of these methods, we again collected tubulin structure (on U2OS cells), but with a different excitation wavelength. Here, the cell is illuminated by light with a wavelength of 642 nm instead of 488 nm (the latter used for the images in the training set). The different wavelength also changes the spatial frequency of the SIM patterns. This cell structure with unseen illumination properties is then propagated through the pretrained models of both SR-REDSIM and RED-fairSIM. The resulting SR-SIM image shows that RED-fairSIM is more robust against changed microscope settings than SR-REDSIM. Scale bar: 4  $\mu\text{m}$ .

further cross-checks should be carried out. As an initial result, RED-fairSIM seems to generalize well to different SIM pattern settings.

### 1. U-Net-fairSIM

In the pipeline proposed for RED-fairSIM, we replaced the RED-Net by U-Net and analyzed the resulting super-resolution images. We named this approach as U-Net-fairSIM. Figure 3 (column 3) shows that U-Net-fairSIM also produces better results as compared to the noisy, reference, and SR-REDSIM images. However, it does not surpass RED-fairSIM. Similarly, the U-Net-fairSIM approach outperforms all other counterparts except for RED-fairSIM concerning the PSNR and SSIM values in Table 1. Comparing RED-fairSIM and U-Net-fairSIM directly, as in Fig. 5, the cell structures reconstructed by RED-fairSIM are smoother. Furthermore, they are more faithful when taking the reference as “gold standard” into account. For these reasons, we have focused the presentation in this paper on RED-fairSIM.

### 2. Data Augmentation

During all of the reported experiments, we did not carry out any preprocessing on the input or output images such as down-sampling or cropping, as discussed in Section 2.C. Nonetheless, we tested whether or not image augmentation yields any significant improvement in the results. For this purpose, each image was rotated by an angle of 180° and added in this form to the training set. This increased the amount of training data from 2025 to 4050 images. The mean PSNR and SSIM values after training the methods with the augmentation approach are reported in Table 2. While the image augmentation did not

**Table 2. Mean PSNR and SSIM Values along with Standard Deviations (STD) for Noise Level 4 after Applying Data Augmentation (Computed on 500 Test Images)**

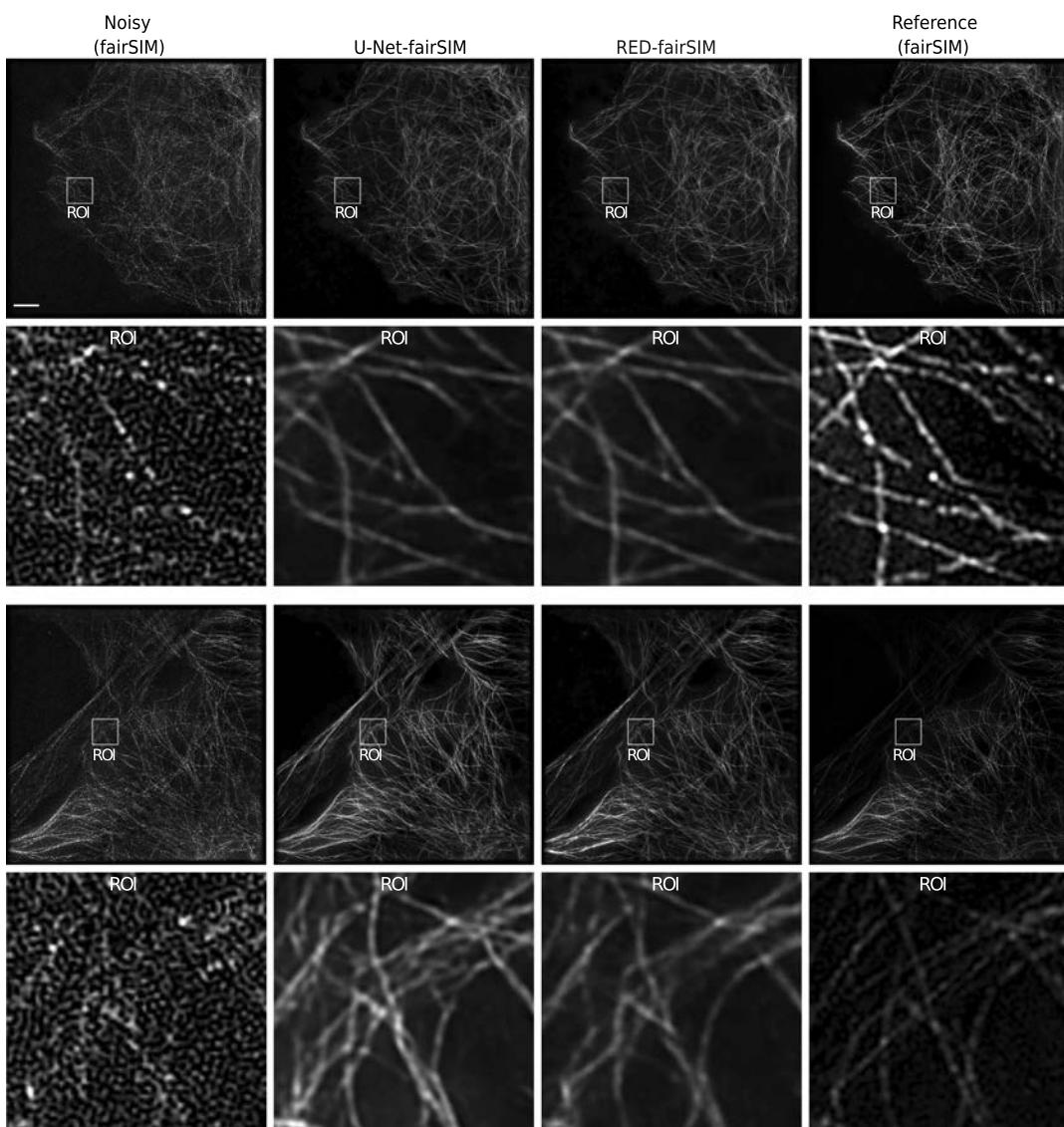
	PSNR(STD)	SSIM(STD)
SR-REDSIM	26.37(2.07)	0.66(0.10)
U-Net-fairSIM	28.05(2.10)	0.71(0.09)
RED-fairSIM	28.09(2.04)	0.72(0.09)

provide a noticeable advantage overall and did not change the performance-wise ordering of the proposed methods, and instead doubled the training time and increased the preprocessing effort, we decided to focus this paper on the results without image augmentation.

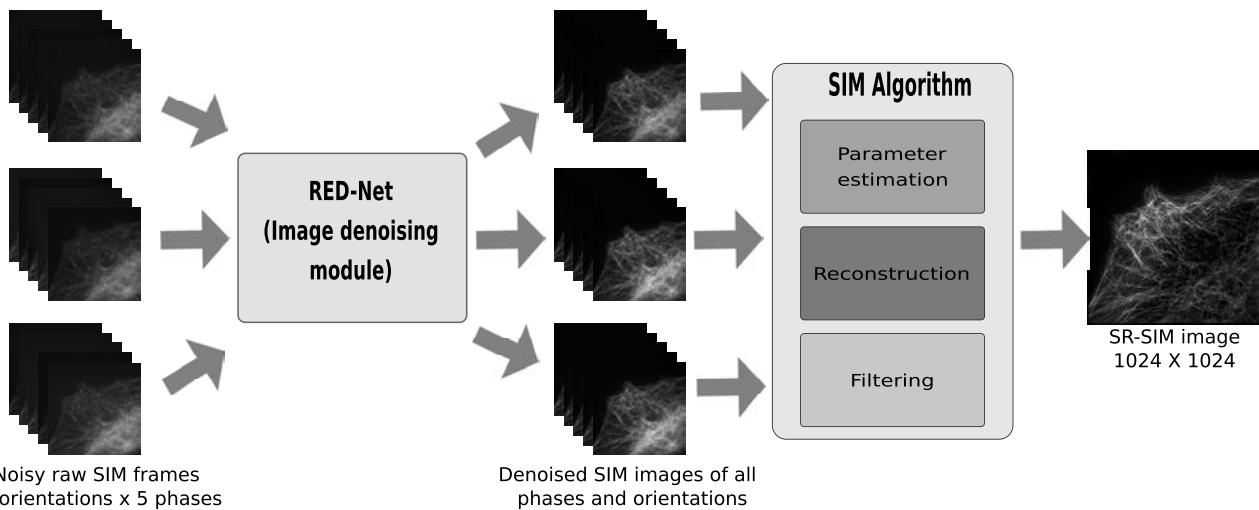
### C. Alternative Approaches

#### 1. preRED-fairSIM

Similarly, we also tried to generate high-quality super-resolution SR-SIM images by another method called preRED-fairSIM. The pipeline of preRED-fairSIM is shown in Fig. 6. However, preRED-fairSIM failed to deliver usable results at the end. In preRED-fairSIM, each noisy SIM image from a different phase and orientation is denoised separately and then the whole stack of all 15 denoised images is propagated into the fairSIM algorithm to reconstruct a final super-resolution image. In the preRED-fairSIM approach, we trained a 30-layer RED-Net for one selected phase and orientation and then performed transfer learning [33] (which implies in our scenario no retraining and no changes in the network weights) and fine tuning [34] (which implies adaptation



**Fig. 5.** Reconstruction of SR-SIM images of two different test samples with the fairSIM, U-Net-fairSIM, and RED-fairSIM methods. Each column represents the results of the corresponding method. The first and third rows show the resulting SR-SIM images, whereas the second and fourth rows contain the extracted enlarged ROIs from the full-size images in the rows directly above. The cell structures reconstructed by RED-fairSIM are smoother compared to the U-Net-fairSIM and fairSIM (reference) cell structures. Furthermore, they are more faithful than U-Net-fairSIM when taking the reference as the “gold standard” into account. Scale bar: 4  $\mu\text{m}$ .



**Fig. 6.** The complete pipeline of the preRED-fairSIM method. In this pipeline, the raw SIM images [512 × 512 (width × height)] of all phases and orientations are denoised separately with the RED-Net architecture. The complete architecture of RED-Net is shown in Fig. 2. The denoised SIM images of each phase and orientation are then propagated into the fairSIM software in the form of a stack (15 frames) to reconstruct the super-resolution SIM image.

of a subset of the network weights) to other phases and orientations.

The results of transfer learning and fine-tuning are quite promising in terms of the achieved SSIM and PSNR values. The model trained with data of phase 0 and orientation 0 has a mean PSNR = 33.22 dB and a mean SSIM = 0.90 on the test data. Averaged over all other combinations of phase and orientation, transfer learning yields a PSNR = 31.03 dB and an SSIM = 0.89 during testing. After fine-tuning the first and last five layers of the pretrained model, the mean values over all combinations of phase and orientation amount to a PSNR = 33.32 dB and an SSIM = 0.89. These numbers suggest that the denoising of the raw SIM images works well for all orientation phases.

The empirical results of this approach on the image level are shown in Fig. 7 (with fine-tuning applied). They also prove that the raw SIM images are well denoised in the first step of this method. However, fairSIM fails to reconstruct super-resolution images of sufficient quality from the denoised raw images. The resulting reconstructed images contain some additional new artifacts. These artifacts can likely be traced to higher harmonics introduced by the RED-Net in the preRED denoising step, which become very clear in the Fourier power spectrum of the denoised images (see Fig. 7), and then appear similarly as artifacts in the Fourier spectrum of the fully reconstructed image. This is to be expected, as the fairSIM method works in the frequency domain, and highly relies on the precise phases, orientations, and harmonics of the SIM pattern, which the denoising step obviously breaks.

## 2. Hessian SIM

Hessian SIM [10] is a conventional, frequency-space-based filtering approach, tailored specifically to reduce noise in the band-limited signal of methods such as structured illumination microscopy. Compared to the other methods presented here and compared against, Hessian SIM aims to reduce noise by

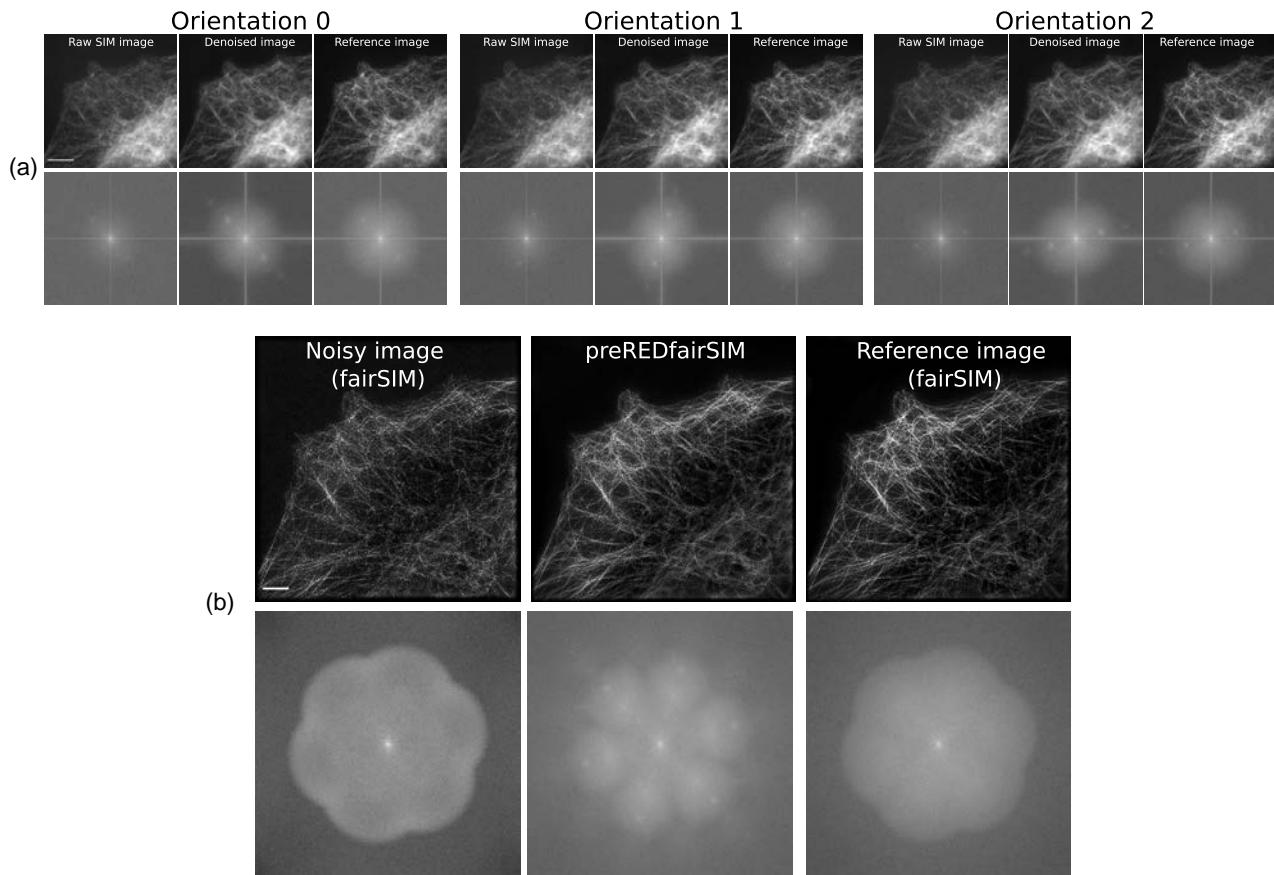
taking into account both the spatial and temporal frequency distribution in the signal. Thus, while the spatial filtering it offers can be applied to single images, it is most successful when applied to time-lapse data with high temporal sampling. This biases the comparison somewhat in favor of Hessian SIM, because applying the algorithm to the time-lapse data as needed for full performance allows it to see more data than the other algorithms, which can only be applied to a single image. Despite using time-lapse data, the Hessian SIM algorithm does not provide good results for noise level 4, as shown in column 4 of Fig. 8, compared to Red-fairSIM and SR-REDSIM. However, the performance of Hessian SIM was reasonable for the data of lower noise levels.

## 3. BM3D

BM3D is a conventional state-of-the-art image denoising method from the field of computer vision [35]. During this work, we also used BM3D to denoise the noisy super-resolution images reconstructed by fairSIM. BM3D is able to remove the noise successfully from reconstructed noisy SIM samples, as shown in column 3 of Fig. 8, but fails to recover lost information.

## D. SIM Reconstruction at Varying Noise Levels

The raw SIM data in this study was collected as a time-lapse of a fixed sample undergoing photobleaching. Thus, data was collected at different noise levels, which can be assembled into five noise level groups, with noise level 0 representing the lowest and noise level 4 representing the highest level of noise found in the data. As we have previously discussed, the models were trained with input from the highest noise level. Therefore, we investigated whether these pretrained models will also be useful for the SIM images of other lower noise levels. To verify this, we considered only the two best methods from our work: SR-REDSIM and RED-fairSIM are used to evaluate the raw SIM images at different levels. No fine-tuning [34] or transfer learning [33] was performed on these pretrained models.



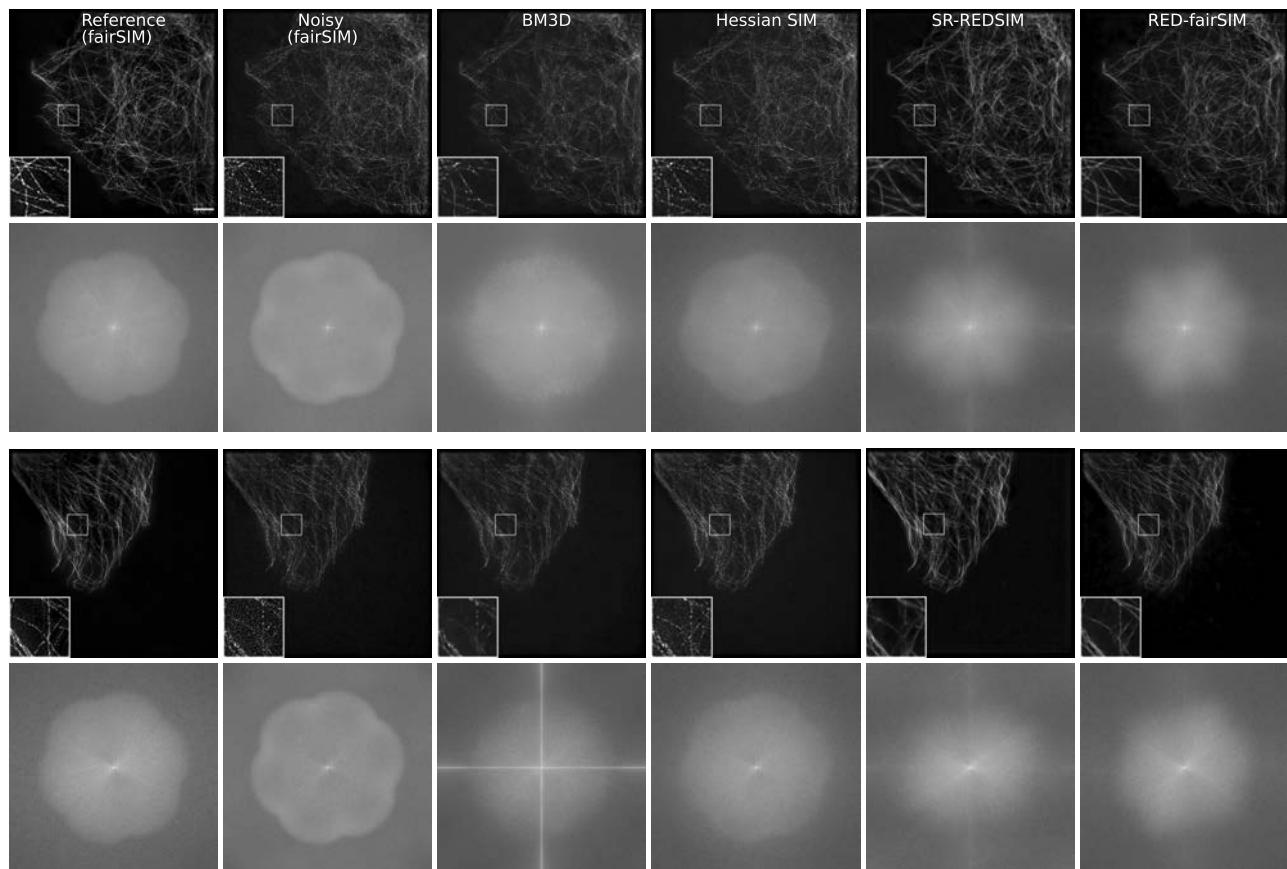
**Fig. 7.** Results for preRED-fairSIM. (a) Three blocks of images where each block consists of six images. The first block depicts the images from phase 0 and orientation 0, the second block from phase 1 and orientation 1, and the third block from phase 2 and orientation 2. The left image in the first row of each block represents a noisy raw SIM image from noise level 4. The second image in each block is the denoised version, whereas the reference image (rightmost in each block) is the ground truth. The Fourier spectra of the images are shown below each image. The dimensions of each image in these blocks are  $512 \times 512$  (width  $\times$  height). Scale bar: 8  $\mu\text{m}$ . (b) Three images reconstructed by fairSIM along with their Fourier spectrum directly below. The noisy image (left) is reconstructed using 15 noisy raw SIM images and the parameter fit summary is: resolution improvement is ( $x:1.90, y:1.90, z:1.90$ ), and the modulation estimation is ( $x:0.310, y:0.341, z:0.332$ ) with the assessment as “weak.” Similarly, the preRED-fairSIM image (middle) is generated using 15 denoised SIM images and the parameter fit summary is: modulation estimation is ( $x:0.310, y:0.341, z:0.332$ ) and the assessment “weak”; however, there is no improvement in the resolution. The reference image (right) is reconstructed from the raw SIM images with the highest SNR and the parameter fit summary is: resolution improvement is ( $x:1.90, y:1.90, z:1.90$ ), and the modulation estimation is ( $x:0.558, y:0.580, z:0.578$ ) with an assessment of “usable.” The Fourier spectrum of preRED-fairSIM shows additional artifacts (white spots) that do not exist in the Fourier spectrum of the reference or the noisy output. Scale bar: 4  $\mu\text{m}$ .

Figure 9 shows the results of this attempt. In this figure, one specific sample was captured at different noise levels. If we further examine the ROIs of all the super-resolution images, it can be seen that both methods show high-quality, super-resolved images at all of the five different noise levels with a slight degradation toward higher noise levels regarding smoothness and clarity of the cell structures. Furthermore, it is again noticeable that the results of RED-fairSIM are overall visually more appealing compared to the other methods.

In addition to visual inspection, quantitative results are given in Table 1, which contains the mean PSNR and SSIM values of 500 test inputs from each noise level. Here, U-Net-fairSIM is included in the comparison. Both RED-fairSIM and U-Net-fairSIM show a gradual decrease in PSNR and SSIM values from noise level 1 (weak noise) to noise level 4 (strong noise). SR-REDSIM performs similarly, but with

noticeably smaller PSNR and SSIM values. The results of fairSIM without denoising deteriorate quickly when moving to higher noise levels. The most important takeaway from Table 1 and Fig. 9 is that the networks—although trained for a specific high noise level—generalize well to conditions with a better SNR.

The reconstruction of the lower noise level 0 through SR-REDSIM and RED-fairSIM highlights a second use case. At this noise level, the SNR of the raw frames is high enough to provide the reference data sets, which, as discussed before, are of high quality, but still feature some SIM reconstruction artifacts. Those artifacts are successfully removed by both SR-REDSIM and RED-fairSIM. A reasonable assumption is that, like noise, reconstruction artifacts are random enough in nature, so they are not picked up by the network during training, and thus cannot be reproduced. This effect is well



**Fig. 8.** The reconstructed SR-SIM images of two test samples with different methods. The Fourier spectrum of each SR-SIM image is shown directly below. Each image contains an enlarged ROI at the lower-left bottom. The analysis of ROI of all the methods clearly shows that the results of RED-fairSIM (sixth column) are smoother and more faithful compared to all other methods. Similarly, the Fourier spectra of the RED-fairSIM do not show any additional artifacts in the Fourier space. The SR-SIM images and ROI of SR-REDSIM (fifth column) also show good results; however, there are some artifacts in the high-frequency region of the Fourier spectrum. The ROIs of Hessian SIM (fourth column) do not outperform RED-fairSIM and SR-REDSIM, but do show better results than BM3D (third column). BM3D produces a suppressed cell structure in both of the resultant images. Furthermore, the Fourier spectrum of the BM3D result for the second test sample shows artifacts in both low and high-frequency regions. Scale bar: 4  $\mu\text{m}$ .

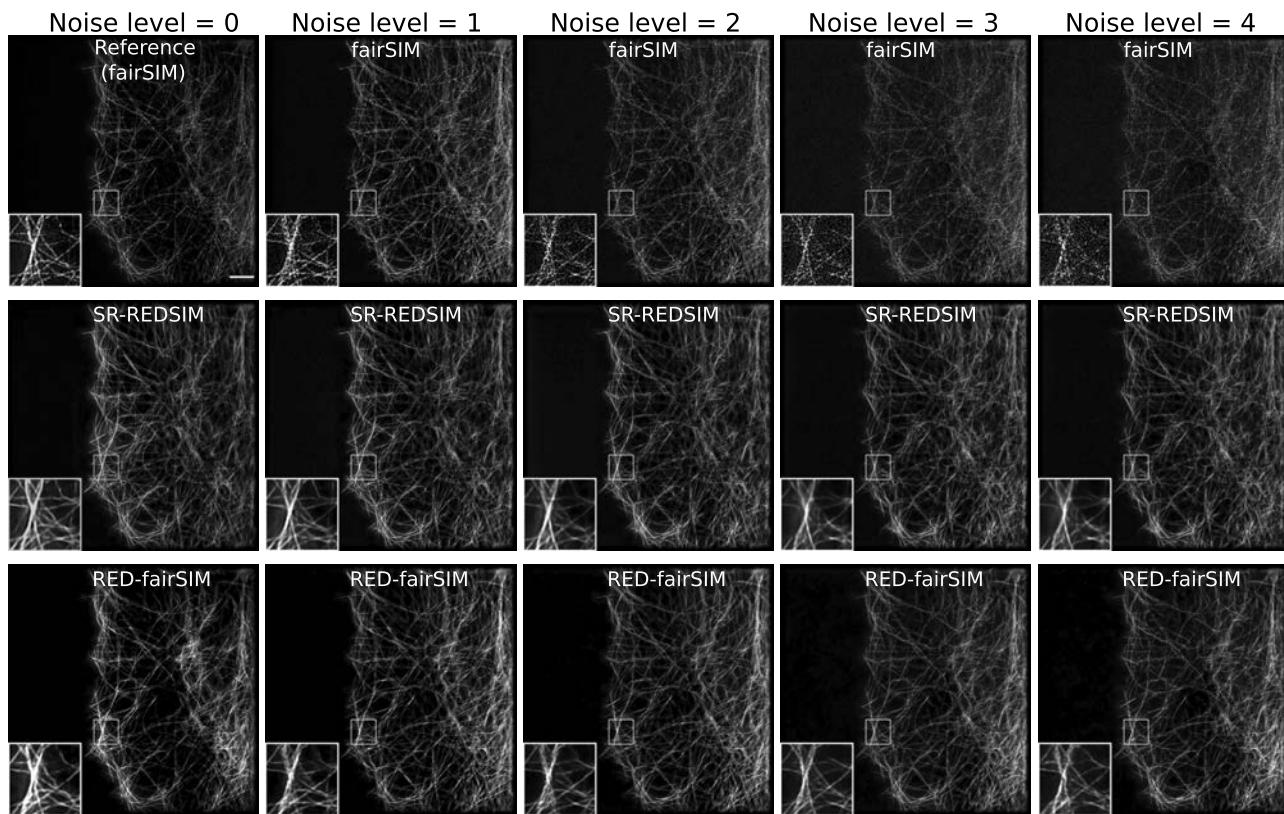
known in applications such as Noise2Noise [36], where the inability of a neural network to learn random (noisy) data is explicitly used for denoising.

#### 4. DISCUSSION

The results of this study provide sufficient evidence that SR-REDSIM and RED-fairSIM can both be employed to denoise and reconstruct high-quality SR-SIM images. In contrast, preRED-fairSIM in its current form is not suitable for this purpose because the output of RED-Net, although successfully denoised, contains additional artifacts noticeable in Fourier space, which spoil the performance of the subsequent classical computational SIM reconstruction. We also investigated the robustness of the successful methods (SR-REDSIM and RED-fairSIM) and showed that high-quality reconstruction of SIM samples is possibly irrespective of the noise level in the raw SIM images. The SR-REDSIM and RED-fairSIM methods outperform their counterparts, as shown in Figs. 3 and 4. Furthermore, these approaches are useful even in the absence

of clean ground-truth data, as we have shown especially for RED-fairSIM where the reference data used for training contains many reconstruction artifacts. We have also shown in column of Fig. 9 that the proposed methods SR-REDSIM and RED-fairSIM can be used to remove the reconstruction artifacts from the reference image after training, so even if high SNR data can be acquired easily, SR-REDSIM and RED-fairSIM still offer an improvement over the classical reconstruction approaches.

A recent study [22] used cycle-consistent generative adversarial networks (CycleGANs) [37] to reconstruct SR-SIM images by using three to nine clean raw SIM images. A CycleGAN contains two generators and two discriminators with multiple losses that are trained in a competitive process. Therefore, CycleGANs are generally very difficult to train. Furthermore, the authors did not address the challenge of denoising. Christensen *et al.* [21] trained deep neural networks by using synthetic data instead of real microscope SIM images to reconstruct SR-SIM images. Although the synthetic data used in their studies for training is unrelated to real micro-



**Fig. 9.** Reconstructed SR-SIM images at different noise levels with the SR-REDSIM and RED-fairSIM methods for a single field of view. Noise level 0 represents the reference image at timestamp 0, noise level 1 comprises the images from timestamps 25–50, noise level 2 from timestamps 75–100, noise level 3 from timestamps 125–150, and noise level 4 from timestamps 175–200. Each full image contains an enlarged ROI in the bottom left. The images reconstructed by fairSIM in the first row show a significant degradation in quality as the noise level increases. In contrast, the results produced by SR-REDSIM and RED-fairSIM in the second and third columns are far less, depending on the noise level. Scale bar: 4  $\mu\text{m}$ .

scopes, they were successful in generating output comparable to computational tools like fairSIM. However, they did not use real noisy microscope data to test the denoising performance of their networks, and their approach was also not completely successful in the case of (simulated) high-level noise. Jin *et al.* [20] used multiple concatenated U-Nets to reconstruct SR-SIM images by using three to 15 raw SIM images. They trained their models on cropped and resized SIM samples and manually discarded tiles with only background information. These preprocessing steps are time-consuming, and the training of two adjacent U-Net models is also computationally expensive.

Our proposed methods use raw SIM images in their original size, which does not involve any major preprocessing steps. The amount of training data used, about 100 fields of view for training and test data together, is also small enough that specific training, capturing both a given instrument and a specific biological structure of interest, should often be feasible. While SR-REDSIM has similarities to other proposed end-to-end deep learning approaches for SIM [20–22], to the best of our knowledge, RED-fairSIM is a completely novel deep learning approach for SIM which is, as our data shows, superior to SR-REDSIM.

While both SR-REDSIM and RED-fairSIM provide high-quality reconstruction, an obvious difference between them is

their ability to generalize to different SIM imaging settings. As an initial test, we varied the spatial frequency of the SIM pattern [using a 642 nm (instead of 488 nm) excitation light], which commonly happens when designing experiments and choosing dyes, as shown in Fig. 4. We then performed reconstruction with RED-fairSIM and SR-REDSIM, both trained on the original 488 nm data. Here, the RED-fairSIM approach, where the change in spatial frequency of the pattern is absorbed by the classic reconstruction step, still works very well in suppressing noise and SIM artifacts. SR-REDSIM, on the other hand, where the SIM pattern has been learned by the network, created heavy ghosting artifacts. While further validation and cross-testing are needed, this suggests that RED-fairSIM should be able to generalize to different SIM microscopes, excitation wavelengths, and probably illumination types (three-beam, two-beam, TIRF-SIM), while SR-REDSIM would require retraining whenever larger changes in these parameters occur.

Besides visual impression, the quantitative measures chosen for our comparisons are PSNR and SSIM. As a microscopy technique, obviously spatial resolution estimates would present another desirable parameter. Both Fourier ring correlation (FRC) [38,39] and image decorrelation analysis (IDA) [40] are typically chosen for this task, because they offer a quantitative resolution estimate that is not dependent on manual

measurement of single structural features, which easily introduces bias. However, much care is needed to correctly apply these methods, as they are based on strict assumptions about the input data, regarding the statistically independent and spatially uncorrelated distribution of noise. Even slight correlation, introduced, for example, by camera readout electronics [41], will yield an overestimation of spatial resolution. We have performed FRC analysis and IDA on the datasets presented here, and while some results seem reasonable, others are clearly unphysical. As we cannot assume that the denoised output fulfills the assumptions of FRC or IDA, we would argue against using unmodified FRC or IDA to estimate resolution in these images. We believe further research into robust resolution estimation that works more independently of the data generation process is needed.

We also compared our set of deep-learning based approaches to classical denoising algorithms, with BM3D as a general approach and Hessian SIM as a noise filter tailored to ideally time-lapsed SIM data. Both of the classical algorithms do not reach the denoising performance of the deep learning methods. However, these algorithms do not require any prior knowledge, in the form of training data, of the sample. Thus, in contrast to deep learning methods, they can be applied to arbitrary, unknown structures. They can also serve as a cross-check, if concerns arise that the deep-learning based approaches generate artifacts stemming from their training data.

## 5. CONCLUSION

In this work, we presented two different methods, SR-REDSIM and RED-fairSIM, to reconstruct super-resolution SIM images from raw SIM images with low SNR. We demonstrated that these methods are robust against different noise intensities and do not need any retraining or fine-tuning even if the SNR is varied between the training and application. However, the generalization ability of RED-fairSIM under different SIM imaging conditions (i.e., changed microscope settings) is superior compared to SR-REDSIM. This shows that the combination of fairSIM for reconstruction and RED-Net for denoising is more promising than an end-to-end deep learning approach like SR-REDSIM. Both methods are particularly useful for SIM images with a low SNR since the traditional reconstruction algorithms cannot denoise and generate reconstruction artifacts in the SR-SIM images. Both of our proposed methods can remove these reconstruction artifacts. The overall results also show that our methods outperform other classical denoising methods like BM3D to denoise the noisy SR-SIM images. Furthermore, the proposed methods can potentially be used in the future to handle live-cell SIM imaging data as well as the reconstruction and denoising of SIM images with low SNR from other biological structures.

**Funding.** Deutsche Forschungsgemeinschaft (415832635); H2020 Marie Skłodowska-Curie Actions (752080); Bundesministerium für Bildung und Forschung (01IS18041C).

**Acknowledgment.** The authors would like to thank Dr. Matthias Fricke and David Pelkmann from the Center for Applied Data Science (CfADS) at Bielefeld University of

Applied Sciences for providing access to the GPU compute cluster and Dr. Wolfgang Hübner for providing U2OS cell data with a red fluorescent tubulin stain. We are also thankful to Wadhab Zai el Amri (CfADS) for precursory work on deep learning for the denoising of microscope images. The authors would additionally like to thank Drs. Florian Jug, Carlas Smith, Peter Dedecker, and Wim Vandenberg for fruitful discussions on the application of deep learning to SIM reconstruction. During this work, we used multiple Nvidia Tesla P100 graphics cards. Each network is trained in parallel with two Tesla P100 to reduce the computational time. The machine learning code is written based on the Python libraries *keras* and *tensorflow*, whereas the *fairSIM* code is written in Java. The code and data used during this work can be provided upon request. The generated results are available at <https://doi.org/10.5281/zenodo.4134841>.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

1. L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. Drummen, "Super-resolution microscopy demystified," *Nat. Cell Biol.* **21**, 72–84 (2019).
2. J. Demmerle, C. Innocent, A. J. North, G. Ball, M. Müller, E. Miron, A. Matsuda, I. M. Dobbie, Y. Markaki, and L. Schermelleh, "Strategic and practical guidelines for successful structured illumination microscopy," *Nat. Protoc.* **12**, 988–1010 (2017).
3. R. Heintzmann and T. Huser, "Super-resolution structured illumination microscopy," *Chem. Rev.* **117**, 13890–13908 (2017).
4. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," *J. Microsc.* **198**, 82–87 (2000).
5. L. M. Hirvonen, K. Wicker, O. Mandula, and R. Heintzmann, "Structured illumination microscopy of a living cell," *Eur. Biophys. J.* **38**, 807–812 (2009).
6. P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. Gustafsson, "Super-resolution video microscopy of live cells by structured illumination," *Nat. Methods* **6**, 339–342 (2009).
7. L. Shao, P. Kner, E. H. Rego, and M. G. Gustafsson, "Super-resolution 3D microscopy of live whole cells using structured illumination," *Nat. Methods* **8**, 1044–1046 (2011).
8. L. Gao, L. Shao, C. D. Higgins, J. S. Poulton, M. Peifer, M. W. Davidson, X. Wu, B. Goldstein, and E. Betzig, "Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens," *Cell* **151**, 1370–1385 (2012).
9. R. Fiolka, L. Shao, E. H. Rego, M. W. Davidson, and M. G. Gustafsson, "Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination," *Proc. Natl. Acad. Sci. USA* **109**, 5311–5315 (2012).
10. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, and L. Chen, "Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy," *Nat. Biotechnol.* **36**, 451–459 (2018).
11. A. Markwirth, M. Lachetta, V. Mönkemöller, R. Heintzmann, W. Hübner, T. Huser, and M. Müller, "Video-rate multi-color structured illumination microscopy with simultaneous real-time reconstruction," *Nat. Commun.* **10**, 4315 (2019).
12. A. Sandmeyer, M. Lachetta, H. Sandmeyer, W. Hübner, T. Huser, and M. Müller, "DMD-based super-resolution structured illumination microscopy visualizes live cell dynamics at high speed and low cost," *bioRxiv* 797670 (2019).
13. M. G. Gustafsson, L. Shao, P. M. Carlton, C. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," *Biophys. J.* **94**, 4957–4970 (2008).

14. M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, and T. Huser, "Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ," *Nat. Commun.* **7**, 10980 (2016).
15. A. Lal, C. Shan, and P. Xi, "Structured illumination microscopy image reconstruction algorithm," *IEEE J. Sel. Top. Quantum Electron.* **22**, 50–63 (2016).
16. P. Křížek, T. Lukeš, M. Ovesný, K. Fliegel, and G. M. Hagen, "SIMToolbox: a MATLAB toolbox for structured illumination fluorescence microscopy," *Bioinformatics* **32**, 318–320 (2016).
17. K. Wicker, O. Mandula, G. Best, R. Fiolka, and R. Heintzmann, "Phase optimisation for structured illumination microscopy," *Opt. Express* **21**, 2032–2049 (2013).
18. J. Fan, X. Huang, L. Li, S. Tan, and L. Chen, "A protocol for structured illumination microscopy with minimal reconstruction artifacts," *Biophys. Rep.* **5**, 80–90 (2019).
19. D. P. Hoffman and E. Betzig, "Tiled reconstruction improves structured illumination microscopy," *bioRxiv* 895318 (2020).
20. L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T. C. Elston, Y. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," *Nat. Commun.* **11**, 1934 (2020).
21. C. N. Christensen, E. N. Ward, P. Lio, and C. F. Kaminski, "ML-SIM: a deep neural network for reconstruction of structured illumination microscopy images," *arXiv:2003.11064* (2020).
22. C. Ling, C. Zhang, M. Wang, F. Meng, L. Du, and X. Yuan, "Fast structured illumination microscopy via deep learning," *Photon. Res.* **8**, 1350–1359 (2020).
23. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, and E. W. Myers, "Content-aware image restoration: pushing the limits of fluorescence microscopy," *Nat. Methods* **15**, 1090–1097 (2018).
24. X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Advances in Neural Information Processing Systems* (2016), pp. 2802–2810.
25. D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nat. Methods* **12**, 717–724 (2015).
26. T. Novák, T. Gajdos, J. Sinkó, G. Szabó, and M. Erdélyi, "TestSTORM: versatile simulator software for multimodal super-resolution localization fluorescence microscopy," *Sci. Rep.* **7**, 951 (2017).
27. B. Lim, S. Son, H. Kim, S. Nah, and K. Lee, "Enhanced deep residual networks for single image super-resolution," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2017), pp. 1132–1140.
28. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *European Conference on Computer Vision (ECCV)* (2018), pp. 286–301.
29. C. Karras, M. Smedh, R. Förster, H. Deschout, J. Fernandez-Rodriguez, and R. Heintzmann, "Successful optimization of reconstruction parameters in structured illumination microscopy—a practical guide," *Opt. Commun.* **436**, 69–75 (2019).
30. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *arXiv:1505.04597* (2015).
31. J. F. Abascal, S. Bussod, N. Ducros, S. Si-Mohamed, P. Douek, C. Chappard, and F. Peyrin, "A residual U-Net network with image prior for 3D image denoising," *HAL hal-02500664* (2020).
32. A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *20th International Conference on Pattern Recognition* (2010), pp. 2366–2369.
33. L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (IGI Global, 2010), pp. 242–264.
34. J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," in *56th Annual Meeting of the Association for Computational Linguistics* (2018), pp. 328–339.
35. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.* **16**, 2080–2095 (2007).
36. J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: learning image restoration without clean data," in *35th International Conference on Machine Learning* (2018), pp. 2965–2974.
37. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision* (2017), pp. 2223–2232.
38. M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," *J. Struct. Biol.* **151**, 250–262 (2005).
39. R. P. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, and B. Rieger, "Measuring image resolution in optical nanoscopy," *Nat. Methods* **10**, 557–562 (2013).
40. A. C. Descloix, K. S. Grussmayer, and A. Radenovic, "Parameter-free image resolution estimation based on decorrelation analysis," *Nat. Methods* **16**, 918–924 (2019).
41. R. Van den Eynde, A. Sandmeyer, W. Vandenberg, S. Duwé, W. Hübner, T. Huser, P. Dedecker, and M. Müller, "Quantitative comparison of camera technologies for cost-effective super-resolution optical fluctuation imaging (SOFI)," *J. Phys. Photon.* **1**, 044001 (2019).

# PHOTONICS Research

## Deep learning in nano-photonics: inverse design and beyond

PETER R. WIECHA,<sup>1,\*</sup> ARNAUD ARBOUET,<sup>2,4</sup> CHRISTIAN GIRARD,<sup>2,5</sup> AND OTTO L. MUSKENS<sup>3,6</sup>

<sup>1</sup>LAAS, Université de Toulouse, CNRS, Toulouse, France

<sup>2</sup>CEMES, Université de Toulouse, CNRS, Toulouse, France

<sup>3</sup>Physics and Astronomy, Faculty of Engineering and Physical Sciences, University of Southampton, Southampton, UK

<sup>4</sup>e-mail: arbouet@cemes.fr

<sup>5</sup>e-mail: girard@cemes.fr

<sup>6</sup>e-mail: o.muskens@soton.ac.uk

\*Corresponding author: pwiecha@laas.fr

Received 2 December 2020; revised 11 January 2021; accepted 27 January 2021; posted 29 January 2021 (Doc. ID 415960); published 14 April 2021

Deep learning in the context of nano-photonics is mostly discussed in terms of its potential for inverse design of photonic devices or nano-structures. Many of the recent works on machine-learning inverse design are highly specific, and the drawbacks of the respective approaches are often not immediately clear. In this review we want therefore to provide a critical review on the capabilities of deep learning for inverse design and the progress which has been made so far. We classify the different deep-learning-based inverse design approaches at a higher level as well as by the context of their respective applications and critically discuss their strengths and weaknesses. While a significant part of the community's attention lies on nano-photonics inverse design, deep learning has evolved as a tool for a large variety of applications. The second part of the review will focus therefore on machine learning research in nano-photonics "beyond inverse design." This spans from physics-informed neural networks for tremendous acceleration of photonics simulations, over sparse data reconstruction, imaging and "knowledge discovery" to experimental applications.

Published by The Optical Society under the terms of the Creative Commons Attribution 4.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

<https://doi.org/10.1364/PRJ.415960>

### 1. INTRODUCTION

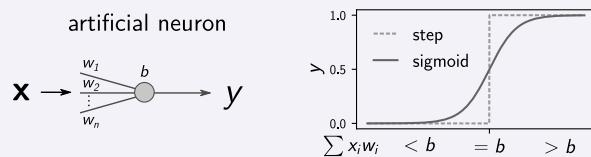
Light-matter interaction at sub-wavelength dimensions can lead to astonishing effects such as localized surface plasmon resonances which concentrate light to deeply sub-wavelength volumes [1], the appearance of optical magnetic resonances in otherwise non-magnetic media [2], the possibility to shape optical near-fields with sub-wavelength structure [3], the emergence of non-linear optical phenomena [4], or strong enhancement of quantum emitter luminescence [5], to name just a few. Those nano-scale optical effects can be exploited for a broad variety of applications, for instance in integrated quantum optics [6], for metamaterials [7], and in this context specifically for metasurfaces like flat lenses [8]. It is, for example, even possible to create all-optical devices which use light to solve integral equations or perform other analog optical computing tasks [9–11].

Still, ever since the advent of nano-optics with the invention of near-field microscopy [12–14] the numerical description of many problems continues to be challenging [15]. An example is the rational design of nano-photonics structures for specific

tasks, which remains a general problem that often involves brute force "forward" calculations or solving inverse scattering problems. Other challenges in nano-optics are related to experimental limitations such as the stochastic nature of single-photon emitters, fluctuating nano-scale force fields such as Brownian motion, and the diffraction limit blocking access to sub-wavelength information. Such effects often complicate the interpretation of nano-optics experiments and require the use of more sophisticated techniques for data analysis, for example, combining data with prior knowledge or sparsity constraints. All these obstacles are about to be pushed significantly further by the emerging computational methods around machine learning. In particular, "deep learning," a sub-field of machine learning which uses complex ANNs with millions of ANs, recently emerged as a versatile and powerful numerical tool [16,17]. Deep learning techniques have proven to be particularly good at the categorization of huge and complex datasets, a task that they perform radically differently compared to classical algorithms. Following a rather "intuitive" approach, ANNs mimic the working principle of biological neurons

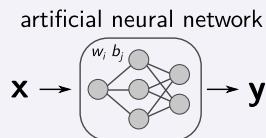
**Box 1. Artificial neurons, neural networks, and their training**

An artificial neuron (AN) is simply a mathematical function which mimics the behavior of a biological neuron.



The step-like behavior of neuronal activation, which starts to fire once a threshold stimulation is exceeded, can be implemented by various non-linear mathematical functions. A popular example is the logistic function (also called "sigmoid"), shown in the above sketch. If the scalar product of an input vector  $\mathbf{x}$  and the neuron-intrinsic weight parameters  $w_i$  are larger than the neuron's bias parameter  $b$ , the output  $y$  is "high" (the artificial neuron fires). Else it is "low."

An artificial neural network (ANN) is composed of several of such ANs, usually arranged in "layers." The output value of a neuron is fed into a succeeding layer of neurons. The final layer is the network output  $\mathbf{y}$ . For instance, in a so-called fully connected ANN, every neuron of one layer is connected to every neuron of the following layer.



Hence, an ANN represents a vectorial function  $f(\mathbf{x}) = \mathbf{y}$  characterized by a large number of parameters  $w_i$  and  $b_j$ .

Training an ANN is done via a numerical minimization of a loss function  $L$ , which describes the error of the

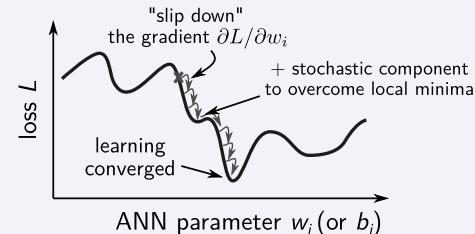
network in predicting samples of the training data. A popular loss function is the mean square error loss (MSE), in particular used for regression tasks:

$$L(w_i, b_j) = \frac{1}{N} \sum_{l=1}^N [y_{\text{train},l} - y_{\text{ANN}}(\mathbf{x}_{\text{train},l})]^2, \quad (1)$$

where  $\mathbf{x}_{\text{train}}$  and  $y_{\text{train}}$  are  $N$  random samples from the training data, and  $y_{\text{ANN}}$  are the network predictions corresponding to  $\mathbf{x}_{\text{train}}$ .  $N$  is called the batch size.

The term "learning" refers to optimizing the parameters  $w_i$  and  $b_j$  describing the ANN, with the goal to minimize the loss  $L$ . A small loss means that the network output approximates well the training data, ideally by learning to "understand" the underlying correlations.  $L$  is numerically minimized by "slipping down" its gradient with respect to the parameters  $w_i$  and  $b_j$ .

training: minimize loss function  $L(w_i, b_j)$



Training on small batches composed of random subsets of  $N$  training samples helps to "jump" out of local minima by adding a stochastic component to the procedure. One of the most common training algorithms is stochastic gradient descent [17].

and the human brain. A brief overview of the basic concepts is given in Box 1.

Research in medicine is often of statistical nature, for which data-driven analysis methods such as deep learning are particularly interesting. Consequently, one of the first scientific fields to which deep learning methods have been extensively applied was medical research. In medical diagnostics, especially medical imaging such as radiology, the use of machine learning techniques for analysis and interpretation has literally exploded in the recent past, which has led to extraordinary successes with diagnostic classification accuracies often far beyond human performance [18,19].

In nano-optics and photonics, machine learning started to emerge a little later, but recently celebrated some remarkable breakthroughs, enabling the analysis, categorization, and interpretation of data which seemed formerly impossible. While already back in the 1990s simple ANNs had been discussed and used for applications in spectroscopy or for automated instrumental control, for instance, to counteract drifts in microscopy [20], it took two decades before the available computational

power reached a level that deep ANNs with millions or even up to hundreds of billions of free parameters [21] could be successfully trained on formerly unsolved problems. Today, deep learning models have evolved to an extent that they readily outperform humans on specialized tasks such as image recognition [16,22]. This progress was possible especially thanks to the rapid development of massive parallel computing architectures in modern graphics processing units (GPUs), and lately of specific "tensor cores," integrated logic circuits optimized for the mathematical matrix operation tasks required for neural network training. Even all-optical implementations of artificial neural networks have been subject to recent research; however, their performance is still limited by the lack of energy-efficient all-optical non-linear units [23–25].

Several review articles have been published recently, which categorize in great detail the latest developments of deep learning applications in photonics and nano-optics. For an exhaustive overview we therefore invite the reader to consult these articles [26–30]. Also a few thematically more distantly related review articles have been published recently, which we want to

indicate to the interested reader. They cover, for example, conventional inverse design and optimization methods for metasurfaces [31] and nano-photonics [32], but also a few more general reviews on artificial intelligence in nano-technology, photonics, and for light–matter interaction have been published [33–36]. Finally, for the sake of conciseness of this review, we intentionally ignore the vast and very active research field on hardware implementations of artificial neural networks, which includes—but is not limited to—research efforts on photonics platforms [23,37,38].

In this mini-review we focus on selected key results that have recently led to breakthrough advancements in the research on inverse design of photonic nano-structures and metasurfaces. Rather than compiling an exhaustive catalog of every single publication, we provide an overview of milestone concepts for improving deep learning inverse design fidelity, which recently allowed to bring ANNs closer to the performance of conventional optimization methods. We believe that such a summary of concepts is of particular interest for researchers in the field. We dedicate the second part of the review to an overview of original applications of deep learning in nano-photonics beyond structural inverse design. Specifically, we summarize recent developments around physics-informed neural networks in optics, on deep learning for knowledge discovery and explainable machine learning, as well as on applications of ANNs to nano-photonics experiments.

## 2. DEEP-LEARNING-BASED NANO-PHOTONICS INVERSE DESIGN

The first part of this mini-review is dedicated to deep-learning-based inverse design techniques as well as to concepts to improve the inverse design model fidelity. As stated before, we do not aim to provide an exhaustive list of applications. An up-to-date and very complete overview of possible optimization targets can be found, for instance, in the recent reviews by Ma *et al.* [27] or by Jiang *et al.* [29].

### A. “Conventional” Inverse Design Methods

Before the recent rise of deep learning methods, inverse design of nano-photonics structures was often based on intuitive considerations and systematic fine-tuning (see, e.g., Refs. [39,40]). A more systematic alternative was the combination of numerical simulation methods with gradient-based or heuristic optimization algorithms, such as stimulated annealing, topology optimization, and genetic algorithms [32,41–44]. Such methods led to some remarkable success for instance in the optimization of plasmonic optical antennas [45,46], dielectric multi-functional nano-structures [47], and metasurfaces [31,48]. A great advantage of such methods is the possibility to include fabrication constraints or robustness conditions in the optimization procedure [47,49].

However, heuristics coupled to numerical simulation techniques is slow and computationally expensive. Furthermore, for each new optimization target, the parameter space needs to be searched from scratch, implying hundreds to thousands of numerical simulations. The recent advent of data-driven techniques such as deep learning holds promise to accelerate the computation by many orders of magnitude and quite some re-

markable progress has been made in the past few years. One can distinguish two types of approach that have gained traction. The first one replaces the forward simulation in an iterative optimization with an ANN, while the second aims to build an inverse ANN that solves the problem directly. Below we critically discuss the two approaches as well as efforts at improving the quality of results.

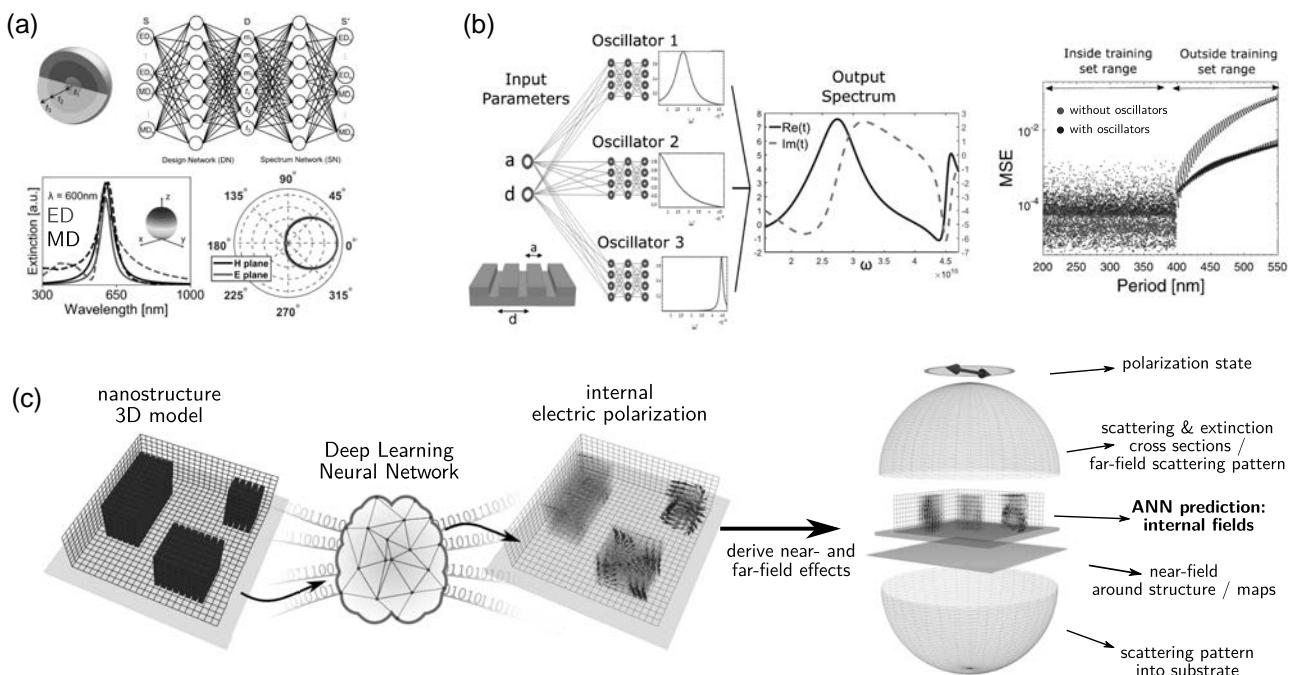
### B. Surrogate-Model-Based Inverse Design

Deep learning models are particularly strong in predicting approximate solutions to direct problems such as the optical response of photonic structures. A possible approach to accelerate inverse design is therefore to use a “forward neural network” as an ultra-fast predictor together with an optimization technique. In such a case the ANN acts as a so-called surrogate model, taking the place of the much slower conventional simulation method.

#### 1. Deep Learning Forward Solver

ANNs have been successfully trained on the prediction of various physical quantities in nano-photonics. Early works have proposed ANNs to create phenomenological models of non-linear optical effects or of optical ionization using experimental training data [50,51]. Recently, the idea has been picked up and it has been shown, for instance, that scattering and extinction spectra can be predicted with high accuracy [52] and also that the phase can be included in the predictions [53], which is important for nano-structures in metasurfaces. The prediction of far-field observables can also be extended to include proximity effects in a dense metasurface, beyond the local phase approximation. The latter has been demonstrated by including the near-field interactions with the nearest neighbor structures in the training data [54]. The prediction of physical effects is not limited to extinction, transmission, or other far-field effects. It has been shown that also near-field effects can be approximated accurately, for instance, around nano-wires of complex shape [55].

While networks that predict an observable such as the scattering cross section perform usually very well within the range of their training data, such models often generalize rather poorly to cases outside the parameter range covered by the training data. The ANN acts then as universal function approximator, but it does not develop a deeper “understanding” of the underlying physics. In order to alleviate this problem, it turned out to be helpful to provide the network with pre-processed data. For instance, instead of training an ANN with pure optical extinction spectra, So *et al.* [56] trained their model using a decomposition in multiple electric and magnetic dipole resonances to predict the optical response of multi-material multi-shell nano-spheres. The approach is illustrated in Fig. 1(a) and has also been used to inverse design multi-shell spheres for Kerker-type directional scattering. Using a metallic grating as a model example, Blanchard-Dionne and Martin demonstrated that a neural network that learns light–matter interaction through a representation as multiple Lorentz oscillators generalizes about an order of magnitude better outside the training data range, compared to a predictor network based on the raw optical spectrum [57] [see Fig. 1(b)]. Instead of predicting specific physical observables such as the extinction cross



**Fig. 1.** Deep-learning-based forward solvers for ultra-fast physics predictions. (a) Simultaneous electric and magnetic dipole resonance prediction and inverse design in multi-layer nano-spheres. Adapted with permission from [56], copyright (2019) American Chemical Society. (b) Nano-optics solver network, which predicts the optical response of a grating based on multiple Lorentz oscillators. As shown in the right panel, the physics-based data representation allows the network to generalize well outside the range of the training data (blue points). Adapted with permission from [57], copyright (2020) Optical Society of America. (c) Internal electric polarization density predictor network. The results can be used in a coupled dipole approximation framework to calculate a large number of secondary near- and far-field effects. Adapted with permission from [58], copyright (2020) American Chemical Society.

section, Wiecha *et al.* demonstrated that a network can learn a discrete dipole approximation of the electric polarization density inside a 3D nano-structure of arbitrary shape [58]. The concept is depicted in Fig. 1(c) and allows accurate derivation of manifold secondary quantities in the near and far fields from a single generalized predictor neural network.

**2. Forward Predictor Networks + Evolutionary Optimization**  
In general, the greatest advantage of deep learning techniques as surrogate models for physics simulations is their tremendous evaluation speed. Once trained, an ANN delivers its prediction within fractions of milliseconds, which is usually orders of magnitude faster than a numerical simulation. Therefore, replacing conventional physics simulations by surrogate ANNs is a natural solution to speed-up the inverse design of photonic nano-structures via global optimization heuristics [59,60]. This concept has recently been applied by several groups to the design of individual photonic nano-structures or metasurfaces [61–66].

However, while the approach can significantly accelerate heuristics-based inverse design, it remains an iterative approach requiring thousands of calls to the surrogate model as well as intermediate computation steps. Furthermore, the surrogate model represents only an approximation to the physical reality, introducing a systematic error. And even worse than that, it cannot be guaranteed that the surrogate model does not contain singular points of totally false solutions [67], to which the optimization algorithm may converge in the worst-case scenario.

Robust implementations therefore require a simulation-based fine-tuning procedure subsequent to the surrogate-based optimization run, which often relativizes the gain in speed [68,69]. The same problem also holds, of course, for the here-after discussed ANN-only inverse design methods.

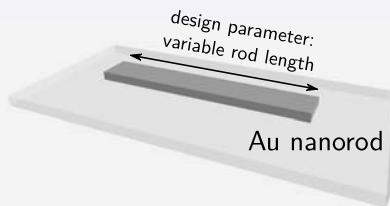
### C. Direct Neural Network Inverse Design

As mentioned above, using forward ANNs as surrogate models for evolutionary optimization is computationally not the most efficient technique and bears the risk of converging to singular points of the surrogate model. In the recent past tremendous efforts have therefore been dedicated to the development of exclusively ANN-based inverse design schemes. The main obstacle which needs to be circumvented is the so-called “one-to-many” problem, which describes the fact that most inverse design problems are ambiguous, and hence several non-unique solutions exist for the same design target. In consequence a naive inversion of the ANN layout usually fails [70], but several solutions have been developed to tackle the one-to-many problem. One possibility is the above-described technique to use a forward network as surrogate model, coupled to a global optimization algorithm. In this section we give a brief overview of pure neural network models to solve non-unique inverse problems. The different concepts are also schematized in Box 2.

A popular type of a stable inverse design network is the so-called tandem network architecture [52,56,70–72]. In a tandem ANN a forward solver network is trained in a first step.

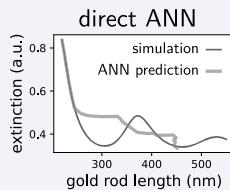
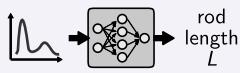
## Box 2. Inverse design: The one-to-many problem

Let us assume a simple toy problem. Under fixed wavelength illumination, we want to tailor the extinction coefficient of a gold nano-rod by varying its length.

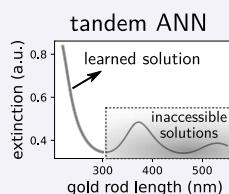
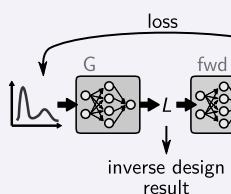


Already this simple problem is ambiguous: several rod lengths can lead to the same extinction, which makes a naive ANN implementation fail in those cases.

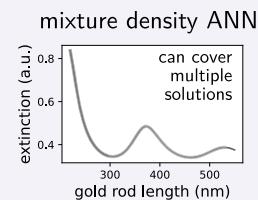
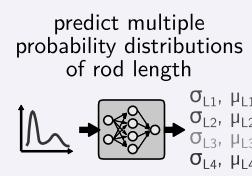
if several solutions exist:  
ANN learns a non-physical  
"average" design parameter



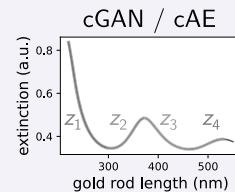
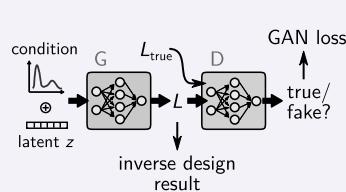
The “tandem neural network” can stabilize the generator (G) via a physics loss, based on a pre-trained forward model (fwd). This approach, however, limits the inverse design to one solution per design target, rendering inaccessible possible multiple solutions to a given problem.



Mixture density networks predict several possible solutions at a time including their respective importance as Gaussian distributions. A disadvantage is that the maximum number of possible simultaneous solutions needs to be known.



Conditional generative adversarial networks (cGANs) or conditional (adversarial) autoencoders add a normally distributed latent vector (usually “z”) to the design target (“condition”), which encodes dynamically multiple possible solutions. Adversarial models furthermore use a trained loss, a so-called discriminator network (D), which tries to distinguish generated (fake) from training samples (true).



The training of the actual inverse design network (the generator) subsequently uses the fixed pre-trained forward model as a physics predictor to evaluate the inverse design output. In consequence, the loss function does not compare ambiguous design layouts but operates in the physics domain (comparing, e.g., the extinction efficiency rather than the design parameters). In this way, different design parameters which lead to a similar physical response no longer confuse the ANN, and all correct solutions to a given design problem yield a positive training feedback.

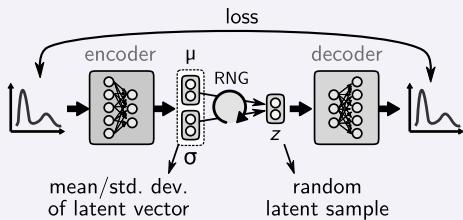
Another model that circumvents the one-to-many problem is the cGAN [68,73–76]. A cGAN takes as input not only the design target but also an additional “latent vector,” which is a normally distributed sequence of random values. The network then learns to use different values of the latent vector to address the distinct non-unique solutions. In addition to the introduction of a latent vector, a further peculiarity of cGANs is their loss function, which is a discriminator network that tries to distinguish generated solutions from real ones, and which is also subject to training. During training, the cGAN loss function

hence evolves together with the ANN, which allows ideally a better convergence. It is worth noting that it is a delicate task to tune the network and training hyperparameters in GANs such that the learning converges. The training of both the generator network and discriminator network needs to evolve in a balanced way for the adversarial loss function to work efficiently.

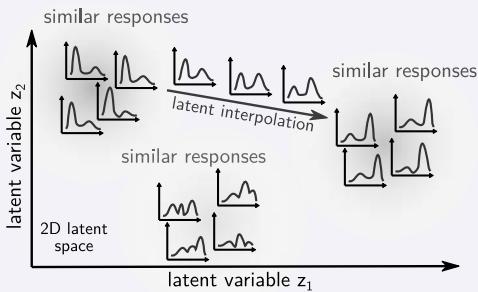
A further type of one-to-many solving networks is conditional adversarial or conditional variational autoencoders [66,77–80]. Those are usually symmetric models that take the physical response as input, which they try to identically reconstruct at their output layer. In a conditional autoencoder, a bottleneck layer is placed in the ANN center. This bottleneck contains the design parameters on the one hand (as it is the case in a tandem network), but on the other hand an additional latent vector is appended to the design parameters. Like in the cGAN, the latent vector can be used by the ANN to address potential multiple solutions. Unlike in the tandem network the forward model is trained simultaneously with the generator. Conditional autoencoders can be seen as a mixture of a tandem

**Box 3. Variational autoencoders and the latent space**

Variational autoencoders (VAEs) learn to compress information in a lower-dimensional latent space, by being trained on a reconstruction task.



In a VAE, forward propagation uses a random number generator (RNG) to draw samples  $z$  with mean value  $\mu$  and standard deviation  $\sigma$ . The random component ensures that the learned latent variables  $z$  follow a normal distribution. However, gradient descent training requires analytical gradients, which cannot be backpropagated through the RNG. This is why a re-parametrization into deterministic layers of  $\mu$  and  $\sigma$  is necessary [81].



By forcing the latent variables on a normal distribution, the trained VAE clusters similar inputs in the latent space. Furthermore, transitions between solutions in latent space are smooth, which allows, for example, interpolation operations.

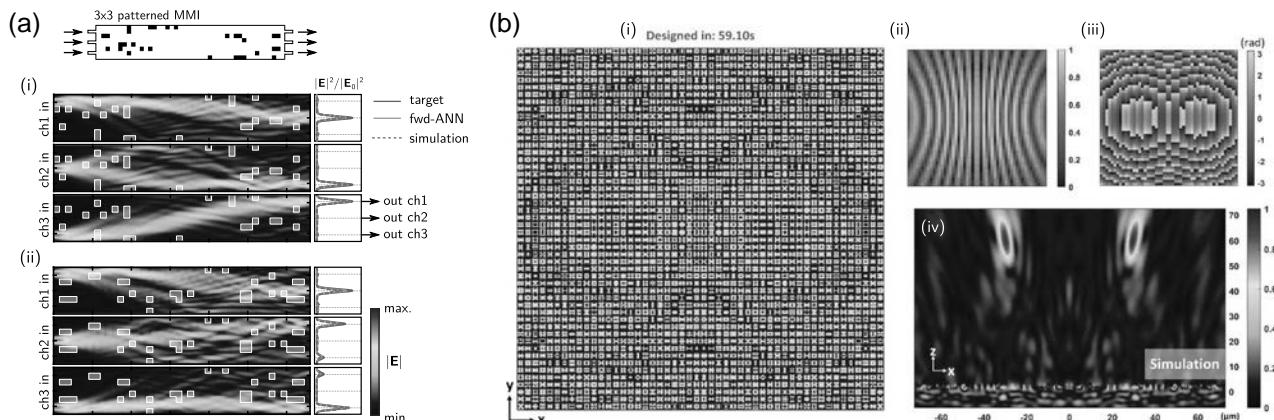
network and a cGAN. For a short explanation of the basic idea behind VAEs and the meaning of the latent space, see also Box 3.

For completeness we want to mention also work on reinforcement learning for iterative design optimization, where the neural network learns to behave as an iterative optimization algorithm. The expectation is that the ANN can adapt its optimization strategy specifically to the given problem and hence outperform conventional heuristic algorithms [82,83].

The above discussed models have been quite successfully used for manifold inverse design problems in nano-photonics. Figure 2(a) shows an example of multi-mode interference devices (MMIs) designed by a tandem ANN. MMIs are large waveguides that support many modes and that can have several inputs and outputs (here  $3 \times 3$ ). The here shown MMIs are patterned with small perturbations in order to obtain specific light-routing properties. The tandem ANN has been trained to design perturbation patterns which produce arbitrary transmission states. This allows, for instance, to define MMI patterns which swap a pair of the  $3 \times 3$  input and output paths, while one of the transmission channels remains constant, as demonstrated in panels (i) and (ii) in Fig. 2(a) [84]. Figure 2(b) shows a metasurface which acts as a flat lens with two focal spots, designed by a variant of a cGAN network [73]. Other examples are the design of chiral plasmonic structures [71,85], dielectric structures [86], multi-shell nano-spheres [56,87], invisibility cloaks [88–90], or metasurface design [91–93].

#### D. Strategies to Improve Neural Network Inverse Design

Data-driven inverse design has the important drawback that the accuracy of the model is first of all limited by the quality of the data and an interpolation error between the data samples is introduced by the ANN. Early works on inverse design therefore reported rather qualitative agreement, but relatively large quantitative inaccuracies. Therefore, in the recent past remarkable



**Fig. 2.** Examples of devices inverse designed by ML algorithms. (a) Encoder-decoder type tandem inverse network used to design perturbation patterns for  $3 \times 3$  MMIs as arbitrary transmission matrix elements. The light routing behavior of the second and the third input channels is interchanged between cases (i) and (ii), while the first input channel keeps routing light to the second output. Adapted with permission from [84], copyright (2021) American Chemical Society. (b) Double-focus flat lens designed by a conditional WGAN inverse network. (i) shows the dielectric metasurface, (ii) the corresponding amplitude, and (iii) the phase mask. (iv) shows a numerical simulation of the field intensity to test the ANN design. Adapted with permission from [73].

efforts have been put in developing methods to improve neural network inverse design. In this section we want to provide an overview over the most successful concepts. In general, two main constituents offer the largest potential for optimization: the training data and the neural network model.

### 1. Improving the Data Quality

As mentioned before, many ANN models do actually generalize relatively poorly to cases outside the parameter range of the training data. They act mainly as generalized function approximators, and hence they interpolate very efficiently to fill the gaps in the training data, while their extrapolation capability remains limited. But also, the interpolation risks may be unsatisfactory if the physical model underlying the training data has sharp features such as high quality factor resonances. If the training data does not contain a sufficient number of such resonant cases, there is a high risk that those features will be very poorly approximated by an ANN.

To tackle this problem, training data can be generated using an optimization algorithm to produce specific responses for the dataset [84,88]. In the case of many free parameters this procedure is time-consuming. Therefore, recently the idea of iterative training data generation has emerged [64,66,84,88, 89,94]. The principal idea is depicted in Fig. 3(a). An initial dataset is generated traditionally via a randomized procedure, on which an inverse design ANN is trained. This network is subsequently used to construct devices based on realistic design targets, but these designs are likely to be rather mediocre as the initial ANN performs relatively poor. Now, the true physical response of these mediocre ANN designs is calculated in another run of numerical simulations, and these samples are appended to the training data. The generator ANN is then trained again on the now extended training data and the generative cycle is repeated. In this way, the neural network can literally learn from its previous mistakes and its performance on the specific design task will significantly improve. Figure 3(a) shows the example of an optical cloak design problem, for which the inverse design accuracy could be improved by more than 1 order of magnitude thanks to iterative training [89]. To visualize the evolution of the training data quality, Fig. 3(b) shows the statistical distributions of resonator quality factors in a fully random dataset of photonic crystal cavities (left) compared to a dataset after one iteration of iterative training (right) [94]. The lack of resonant geometries in the randomly generated dataset is evident. Despite those solutions not being present in the initial dataset, the ANN managed to conjecture a certain amount of resonant cases, improving the training data for the second iteration. By repeating the procedure, the training data increasingly contains resonant geometries, which consequently allows the ANN to inverse design close-to-optimum solutions. Another positive side effect specifically in tandem networks is that iterative training simultaneously improves the accuracy of the forward network [84]. A recent work showed that an even better design performance can be achieved by iteratively increasing the network complexity together with a successive augmentation of the training data, as depicted in Fig. 3(c) [95].

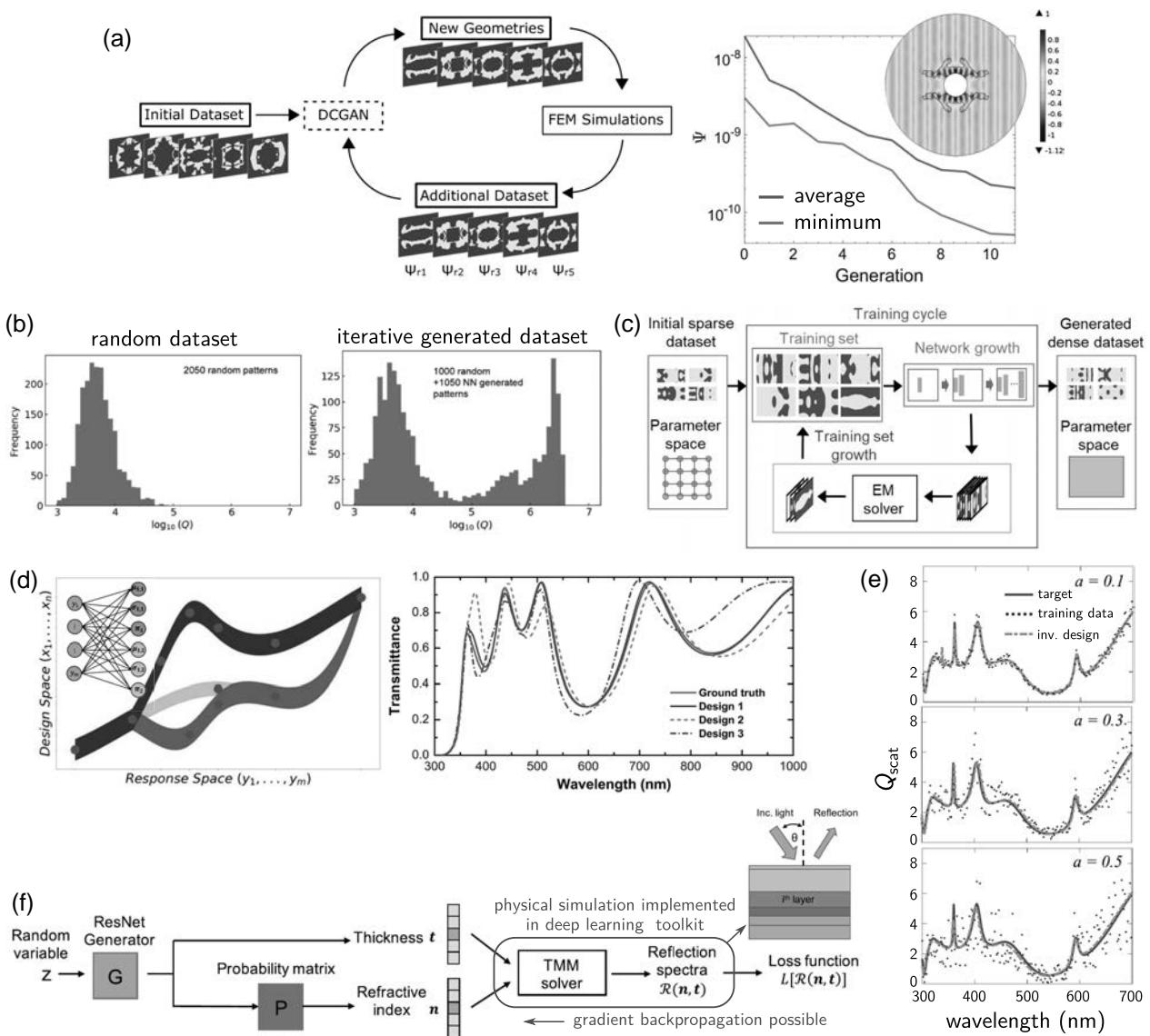
An obvious drawback of iterative procedures is their computational cost. Data generation is usually slow, and the expensive network training needs to be repeated several times on

increasing amounts of training samples. Several suggestions have been made to accelerate the convergence of iterative data generation in order to reduce the number of cycles. For instance, by training several networks, the statistics from multiple predictions can be used to assess the quality and the uncertainty of the ANN output (“wisdom of the many” [96]; see also Box 4). This information can be exploited to choose only the best new solutions for re-simulation and insertion into the expanded training data, which reduces the number of expensive physics simulations [64]. Similarly, an evolutionary optimization algorithm might be coupled to a generative ANN in the iterative cycle to further specialize the training data with regards to the anticipated optimization target [66]. A drawback of such training-data optimization strategies is a risk of over-specializing the network to optimum cases and losing its capability to generalize to arbitrary situations. Therefore, care needs to be taken that the training data remains sufficiently diverse.

### 2. Physics-Model-Based Loss Function

A similar, yet somehow more radical concept is to not use a fixed set of training data at all but instead to implement a loss function based on a physical model within the framework of the machine learning toolkit. Such an approach has been illustrated recently by the example of inverse designing multi-layer thin-film stacks for specific reflection and transmission spectra [97]. As highlighted by a red box on the right in Fig. 3(f), a transfer matrix method (TMM) has been implemented directly in the deep learning toolkit as a loss function. In consequence, error backpropagation is possible through the TMM solver, and the network can be trained without an explicit dataset. The loss function in this so-called “GLOnet” is used to optimize the transmission and reflection spectra of a multi-layer stack with respect to a design target. It is worth mentioning that the GLOnet learns to optimize a single design target, and hence in principle the training of the network takes the place of a conventional global optimization algorithm run (hence its name “GLOnet”). The authors of Ref. [97] claim that the training dynamics allow their GLOnet to ideally adapt its optimization scheme to each problem, resulting in better and faster convergence compared to hard-coded optimizers. The same authors have generalized their concept to a somehow more flexible inverse network called “conditional GLOnet,” using an iterative training scheme instead of a fully differentiable physics loss function. For the training, gradients of the design efficiency are calculated via adjoint simulations and re-injected for back-propagation through the network [98]. The conditional GLOnet is conceptually similar to a Pareto optimization in which a set of optimum solutions for a multi-objective problem is calculated [99]. While the specific solving of a single problem is intentional in Refs. [97,98], as already mentioned before over-specialization is an inherent danger of all iterative data-generation methods.

Another concept to replace the dataset by a direct evaluation of a physics model has been demonstrated for the Helmholtz equation, by developing a loss function which directly evaluates this partial differential equation (PDE). Such an ANN model is called a “physics-informed neural network” (PINN). In the case of a Helmholtz-PINN, the network learns to directly solve the wave equation in the frequency domain. The inverse design



**Fig. 3.** Concepts to improve common shortcomings of inverse design ANNs. (a) Iterative training data generation, in which a network learns from its own errors, here applied to the inverse design of an invisibility cloak device. Adapted from [89], copyright (2021) Optical Society of America. (b) Comparison of the  $Q$ -factors for photonic crystal cavities in a random dataset (left) and in an iteratively generated dataset after the first iteration (right). Adapted from [94], copyright (2019) de Gruyter. (c) Together with the training data, the network complexity can be progressively growing, allowing even better performance by successive learning of smaller features. Reprinted with permission from [95], copyright (2020) American Chemical Society. (d) Mixture density ANN which represents multiple solutions with Gaussian probability distributions to find several non-unique solutions to ambiguous problems. The shown example deals with the spectral design of a multi-layer stack. Adapted with permission from [100], copyright (2020) American Chemical Society. (e) De-noising inverse ANN as robust approach for training on noisy data (noise parameter  $a$  increasing from top to bottom). Adapted from [101], copyright (2019) Optical Society of America. (f) “GLONet”: inverse design ANN using a transfer-matrix model loss for reflectivity and transmission spectra optimization of multi-layer stacks. Adapted from [97], copyright (2020) de Gruyter.

target is then implemented as a boundary condition matching problem [90,102]. As in the GLOnet case, also such a PINN inverse design requires a new training run for each optimization target. PINNs will be discussed in more detail later in this review.

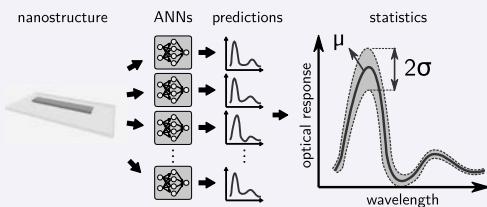
### 3. Sophisticated ANN Models

The second main lever allowing for performance optimization of inverse design ANNs is the neural network model itself.

It has been proven helpful to adopt recent findings in the research on optimum network layout for deep learning. For instance, if applicable the “U-Net” architecture [103] offers much better training convergence and generalization capacity than standard convolutional neural networks—even in cases where its particularly efficient segmentation capabilities are not required [58,104]. Furthermore, so-called residual blocks, or ResNets [22], should be adopted whenever possible. Residual blocks are characterized by their skip connections which avoid

**Box 4. “Wisdom of the many”**

*Wisdom of the many* or also *wisdom of the crowd* denotes the procedure of training multiple neural networks on the same data, each ANN with random initialization. We illustrate the idea by the example of an optical spectrum predictor network.



While this approach adds a significant computational cost (training several networks), the mean  $\mu$  of  $N$  independent predictions provides a  $\sqrt{N}$  times smaller statistical error compared to using a single ANN. Furthermore, the standard deviation  $\sigma$  of multiple predictions can be used to assess the credence of the ANN output.

the vanishing gradient problem, allowing the training of very deep network layouts.

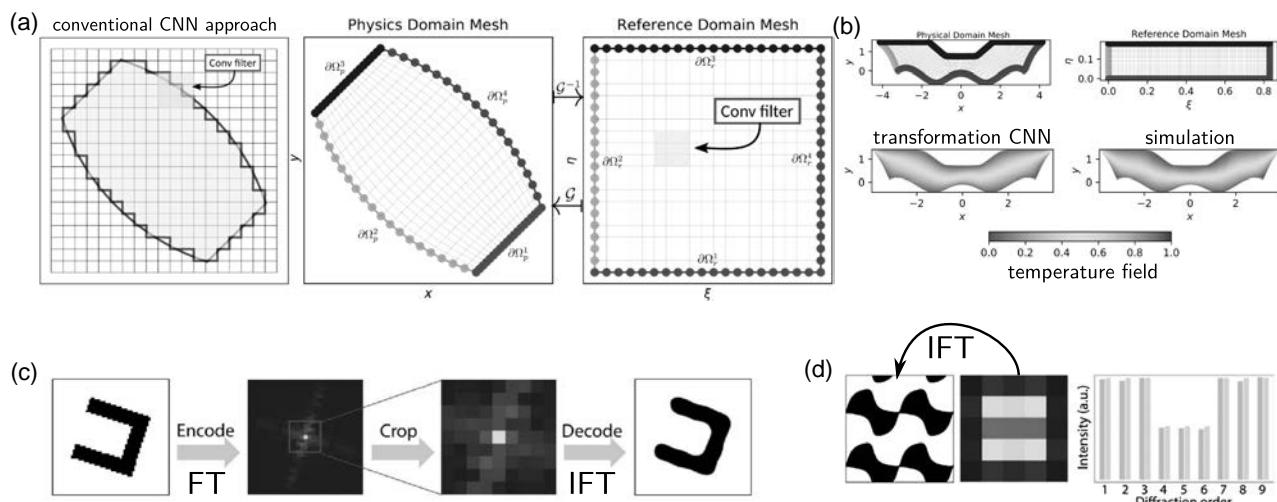
In addition to the application of general “best-practice” ANN design rules, problem-specific tailoring of the network layout can be very favorable for optimum inverse design performance. For instance, to tackle the one-to-many problem, “multi-branch” or “mixture density” ANNs can be applied in addition to the above-named network architectures. The concept is based on representing the design parameters in a “modal” representation as multiple Gaussian distributions,

where each of the Gaussian distributions describes a possible solution to an ambiguous problem (see also Box 2). This concept was proposed some time ago for microwave device inverse design [105,106] and was recently adapted to nano-photonics [100,107] [see also Fig. 3(d)]. The advantage is that the network can in principle deliver all possible solutions together with a weight for their respective priorities. A drawback of the approach is that the approximate number of non-unique solutions needs to be known in advance.

Another recent proposition to optimize inverse networks specifically for noisy situations like in experiments is the implementation of concepts from machine-learning-based image denoising [108]. As shown in Fig. 3(e), Hu *et al.* added artificial noise on training data and could demonstrate that a denoising network-based inverse ANN offers a very robust performance even when trained on very noisy data [101]. This opens promising perspectives for experimental applications.

#### 4. Reformatting the Input Data

Apart from optimizing the network model and generating training data of high quality, the format of the inputs and outputs of a neural network can play a decisive role in whether the ANN manages to “understand” the data or not. An example is illustrated in Fig. 4(a), where a physical problem is to be solved on a non-Cartesian coordinate domain. On 2D problems such as the one here shown, typically convolutional neural networks (CNNs) are most efficient. However, as can be seen in the left-most panel, the imposed discretization on a square mesh is very poor. This holds in particular for the domain borders. Gao *et al.* [109] proposed therefore to apply a transformation of the coordinate system from the physics domain to the CNN reference domain prior training. As illustrated in Fig. 4(b) by the example of solving the heat equation, this additional pre-processing allows to successfully apply ANNs to very complex non-uniform physical domains.



**Fig. 4.** Examples of input data pre-processing for optimized physics domain representation. (a),(b) Deep learning on irregular grids via coordinate transform (a) which is implemented within the deep learning toolkit to allow fast gradient calculations through the coordinate system transformation. (b) The transformation allows to efficiently train networks on complex-shaped physical domains. Adapted with permission from [109], copyright (2020) Elsevier. (c) Data encoding and compression using a topology description based on low-frequency Fourier components, which allows data-efficient treatment of complex shapes, here for example a free-form metagrating. Adapted from [110], copyright (2020) Optical Society of America.

The problem of discretization can also be alleviated by applying a topology encoding procedure, for instance via Fourier transformation [110]. The idea is illustrated in Figs. 4(c) and 4(d). Such encoding can allow not only to describe geometries with odd shapes without restrictions due to discretization, but it allows furthermore to condense the information to a low-dimensional space, which is helpful to reduce ANN complexity and furthermore advantageous in preventing overfitting.

### 5. Other Concepts

Further possibilities to improve the quality of ANN-based inverse design are to use the ANN only as a first step for a rough estimate and apply a conventional iterative approach in a subsequent refinement step. Heuristic optimization algorithms usually benefit strongly from a good initial guess [68]. Another recent proposition is to use a forward neural network purely as an ultra-fast physics predictor to construct a huge lookup table [111]. Using a well-trained forward network, a lookup table can be created which covers the entire parameter space at a very fine resolution, impossible to achieve with conventional numerical methods. Appropriate solutions to specific problems can subsequently be searched in this database. Transfer learning has also been recently applied to nano-optics problems to improve ANN performance if only small amounts of data exist [112]. For instance, experimental data is often expensive, but the situation can be improved by training an ANN first on simulated data, and subsequently specializing the pre-trained network via transfer learning on the experimental dataset [113].

### E. Heuristics versus Deep Learning—A Critical Comparison

It is of utmost importance to emphasize that a data-driven inverse design technique can never outperform an iterative method if it is based on the same simulation model used for training data generation. At least not if no time constraint is set for the iterative optimization. Well-trained and optimized data-driven ANNs usually produce errors of the order of a few percent [55,58]. Furthermore, it is virtually impossible to completely suppress outliers in the network predictions [67]. At the singular points the error of the ANN can be orders of magnitude higher. It is thus a delicate task to assess whether a prediction is valid or rather the result of a singularity in the ANN.

While recently some sophisticated training techniques were presented that are capable to train ANNs for performances similar to conventional inverse optimization, they are either still considerably constrained or the high accuracy has a severe impact on the computational cost. Examples are physics-loss based inverse ANNs or networks based on progressive-complexity training schemes [95,97]. The model described in Ref. [97], for example, is constrained to a simple transfer-matrix description of a multi-layer system as well as to the inverse design of a single optimization target.

The fact that ANNs always introduce an additional error is inherent to the data-driven nature of machine learning (ML), which implies that an ML model can never outperform the accuracy of the simulations used to create the dataset or the model defining the training loss. On the other hand, once trained ANN techniques can offer extreme speed-up of the inverse design,

generally many orders of magnitude faster than iterative approaches based on numerical simulations, it is not unusual that milliseconds stand against hours or even days. This is a marvelous advantage and often well worth it to accept the reduced accuracy of ANN-based techniques. In daily applications a few percent error might actually not even matter too much, in particular when compared to the typical magnitude of inaccuracies in fabrication.

On the other hand, concerning the inverse design speed it is important to remember that the ultra-fast predictions require a fully trained neural network. This implies the computationally highly demanding data generation as well as the very expensive training of the ANN. In many situations, conventional global optimization is in sum actually computationally cheaper. In conclusion, deep-learning-based inverse design is mainly interesting for applications which require a large number of repetitions of similar design tasks, or that rely on ultimate speed for the design generation.

## 3. BEYOND INVERSE DESIGN

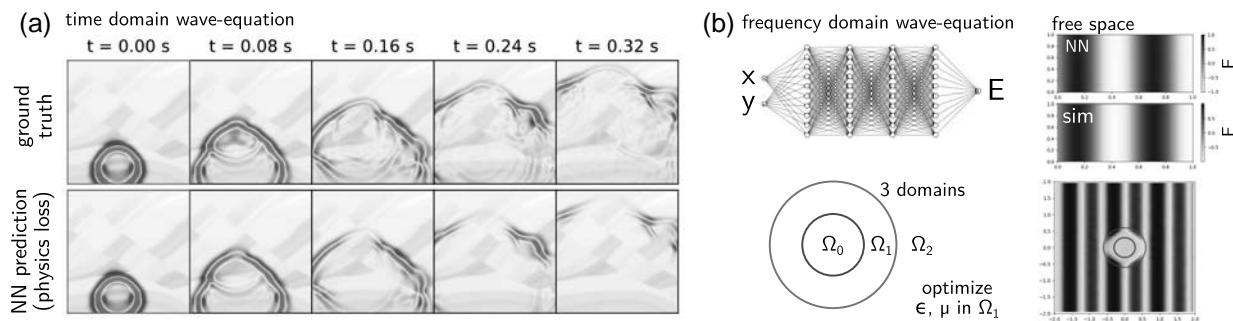
The second part of this review is dedicated to applications of deep learning in nano-photonics “beyond inverse design.” We give an overview on physics-informed neural networks; we present recent work on ANNs for physics interpretation and knowledge discovery as well as experimental applications.

### A. Physics-Informed Neural Networks: Solving PDEs

Most machine learning applications in physics aim to predict derived observables such as transmittance or extinction cross sections. In contrast, the idea of PINNs is to train an ANN to directly predict the solution of a PDE. While this would be also possible using a dataset of pre-calculated solutions, the particularity of PINNs is that instead of using a loss function for data comparison like MSE, the PINN-loss implements an explicit evaluation of the PDE. In consequence, no pre-calculated training data is required for training. For the PINN-loss, the PDE derivatives of the ANN-predicted observables are directly implemented in the respective deep learning toolkit. Thus, the PINN-loss can be seen as a consistency check for the predicted solution. Because modern deep learning toolkits offer powerful automatic differentiation functionalities, error backpropagation through the PINN-loss remains possible and the ANN can be efficiently trained without data.

This concept was first proposed in 2019 by Raissi *et al.* [114] and has since then attracted a great deal of attention across countless research communities in physics, such as fluid mechanics [114,115], thermodynamics [109], and geophysics [116]. Compared to data-based ANNs, the accuracy of PINNs is in general significantly higher. On the other hand, because PINNs evaluate the underlying PDE “point by point,” they are usually slower than conventional data-based models. Since the latter work on physical observables, it is easier to predict higher-dimensional data structures at a time, making better use of the massive parallel computing architectures of modern GPUs. Nevertheless, PINNs are usually orders of magnitude faster than numerical PDE solvers.

Applications to nano-photonics are still scarce. Recently Moseley *et al.* demonstrated that PINNs are capable of



**Fig. 5.** Physics-informed neural networks (PINNs) for nano-optics. (a) PINN for solving the wave equation in the time domain. Adapted with permission from [116]. (b) Top: solving the Helmholtz equation (frequency domain); bottom: using the PINN for inverse design of the permittivity distribution in domain  $\Omega_1$  for an invisibility cloak application. Adapted with permission from [102], copyright (2020) IEEE.

accurately solving the wave equation in the time domain [116]. An example is shown in Fig. 5(a), demonstrating seismic wave propagation through an inhomogeneous medium at successive snapshots in time. As can be seen, the PINN is capable of predicting the evolution of the wave propagation even in a complex environment. While Ref. [116] treats shock waves in geophysics, the problem is conceptually identical to the wave equation in electrodynamics.

Depicted in Fig. 5(b), Fang and Zhan recently demonstrated that a PINN can accurately solve the Helmholtz equation, describing wave propagation in the frequency domain [102]. They found that sinusoidal activation functions are the most adequate choice to solve a differential equation with time-harmonic solutions. By formulating the inverse design as a boundary condition matching problem, it was possible to use the Helmholtz-PINN for the design of an optical cloak, as illustrated in the bottom of Fig. 5(b). A similar frequency-domain PINN has been proposed for the homogenization of optical metamaterials [90]. A disadvantage of PINNs is that the environment needs to be defined at the training stage and hence a new network needs to be trained if the boundary conditions change. Each PINN-based inverse design therefore involves a new training procedure, comparable with conventional iterative techniques, which is evidently much slower than “direct” inverse ANN models. Conceptually related to PINNs is also the so-called “GLOnet,” which is discussed in more detail above [97] [see also Fig. 3(f)].

## B. Interpretation of Physical Properties

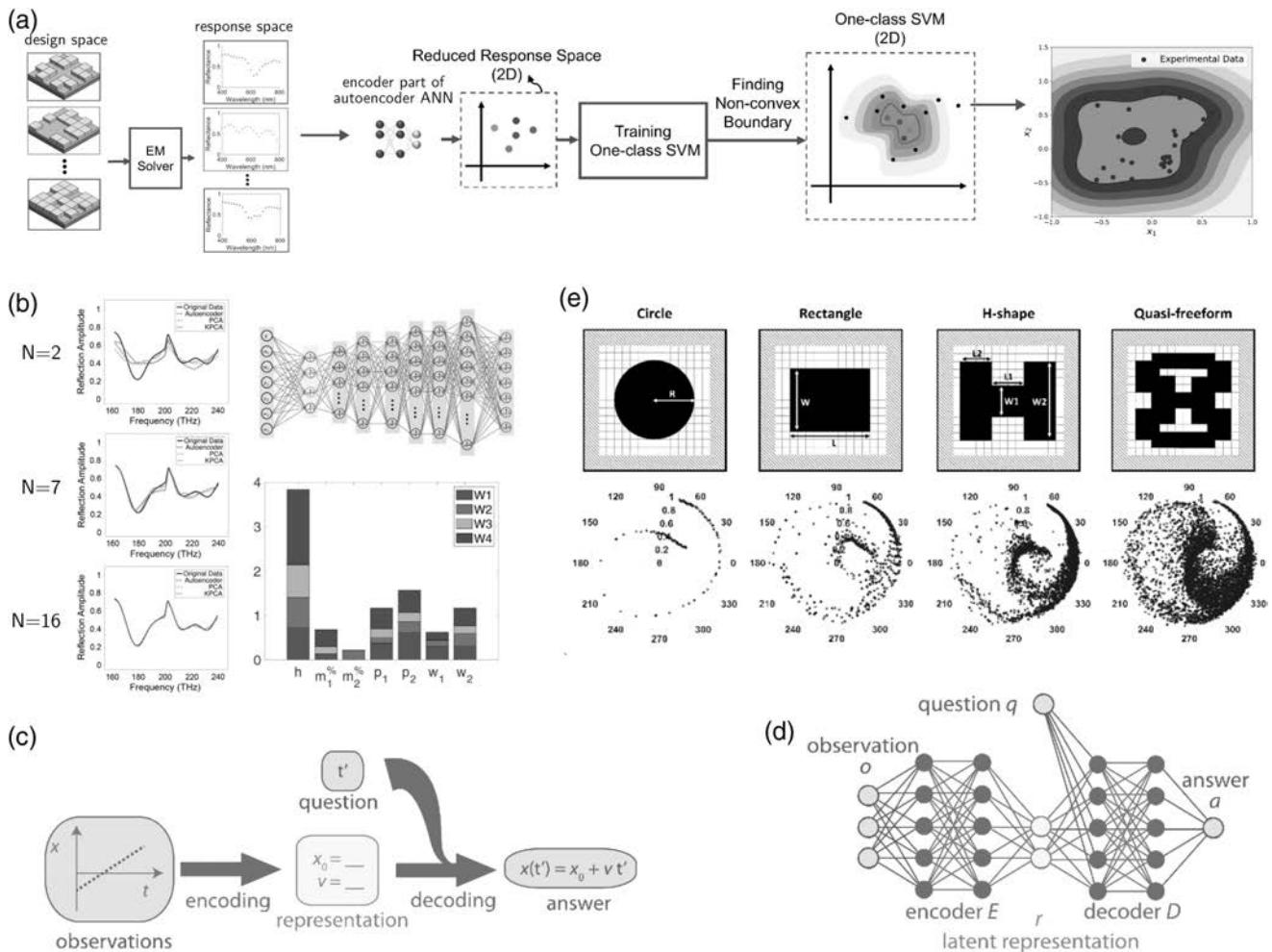
In this section we will review recent approaches to extract information and correlations from deep learning models in order to reveal physical insights.

There is on the one hand the possibility to use deep learning models for dimensionality reduction. Figure 6(a) shows a work of Kiarashinejad *et al.* in which the number of discrete values in reflectance spectra from a set of electrodynamical simulations is reduced from 200 to 2 via an unsupervised autoencoder ANN. In a second step, the non-convex hull of the compressed responses is calculated, which represents the region in the 2D compressed space containing all encoded points. This region allows us to assess the range of accessible physical responses within the allowed design parameters, and the method is hence helpful to identify the physical limitations of specific

nano-structure models. Note that the full physical response of any point in the reduced dimensionality space can be reconstructed using the decoder part of the autoencoder, also those points that were not present in the training data. This means that feasible and non-feasible responses can be analyzed in the original response space (under the assumption that the neural network generalizes well to out-of-training situations). Note that autoencoders are unsupervised ANN models, which are known to require relatively few data for training. This facilitates the application of the technique to experimental data.

In a similar approach, the impact of variations of individual design parameters on the latent space can be studied. Those parameters whose variations have large (respectively little) impact on the latent space contribute strongly (respectively weakly) to the optical response [117–119]. The latent space is indicated by yellow highlighted neurons in Fig. 6(b), top right. The impact of physical parameters on these weights is illustrated in the bottom right of Fig. 6(b). By varying the size of the bottleneck (i.e., reducing the latent space dimension), it is furthermore possible to extract something like the number of principal components of the response, as shown in the left column of Fig. 6(b). Iten *et al.* [120] extended the encoder-decoder ANN for interpretable physics via an approach inspired by humans’ interpretation and modeling of physical observations. The concept is depicted in Fig. 6(c), where the motion of a mass is observed as a function of time  $x(t)$ . To implement this concept in an ANN the authors append a condition to the latent vector at the bottleneck of an encoder-decoder ANN [see Fig. 6(d)]. This condition is here called a question; the example in Fig. 6(c) uses the time  $t'$  for which the ANN shall predict the position of the moving mass (= the answer). In the context of nano-photonics the question could be an optical spectrum of a nano-structure. The “answer” returned by the ANN might then be the material or the size of the nano-structure, or a wavelength or laser polarization state. This kind of ANN is conceptually very similar to inverse design ANNs (in particular to the cGAN or cAE models), but instead of using it for the design of nano-structures, it is here used to understand causal correlations imposed by the implicit physics in the training data.

A more direct approach to extract physical knowledge from ANNs consists in using the ultra-fast approximation capability



**Fig. 6.** Examples of “knowledge discovery” through machine learning. (a) The feasibility of a physical response by a defined geometric model can be assessed by a dimensionality reduction through an autoencoder neural network and subsequent non-convex hull determination. Adapted from [121], copyright (2019) the authors. (b) Study of the impact of the number of bottleneck neurons  $N$  (left spectra) as well as of nano-structure design variations on the activation of the bottleneck neurons ( $W_1$ – $W_4$  in case  $N = 4$ , yellow neurons in the top right panel). This analysis allows to assess the physical importance of individual design parameters and reveals information about the complexity of the optical response. Adapted with permission from [117], copyright (2019) John Wiley and Sons. (c), (d) By mimicking the human approach of interpreting and modeling physical observations (c), a conditional encoder–decoder network (d) can be used to discover implicit physics concepts from data. Reprinted with permission from [120], copyright (2020) APS. (e) Exploiting the high speed of a physics predictor network permits a systematic analysis of the achievable phase and intensity variations in metasurface constituent design. Adapted from [122], copyright (2020) Optical Society of America.

of deep learning surrogate models. Through a systematic scan of the whole parameter space it is, for example, possible to assess the accessible optical responses with a specific nano-structure model. In this way, accessible phase and intensity values for metasurface elements have been classified systematically by An *et al.* [122]. The logical conclusion of the study was that allowing more complex shapes for the meta-atoms leads to a larger accessible range for the phase and intensity, as depicted in Fig. 6(e). From left to right are shown increasingly complex geometric models (top row) and their accessible scattering phase and intensity range (bottom row).

As already mentioned before, another way to gain insight in physical processes through a machine learning analysis is to use a physical parametrization of the training data, such that the neural network explicitly returns a physical quantity. As shown

in Figs. 1(a) and 1(b), extinction spectra can, for example, be pre-processed in a modal decomposition, such as a superposition of electric and magnetic dipole resonances [56] or as a decomposition in Lorentzian resonance profiles [57]. Once trained, the respective neural networks deliver an explicit interpretation of the predicted spectra.

In another recent work, so-called explainable machine learning has been used to assess the importance of constituent parts of a nano-structure with respect to its optical response, as well as to identify those parts of the structure that contribute only weakly to the light–matter interaction [123]. Such information is important for the design of fabrication-robust nano-structures, but also for applications in which sub-constituents of high impact on the nano-structure’s optical response need to be identified, e.g., for switchable optical antennas. Another

recent work proposes interpretable machine learning models like decision trees and random forests to understand the physical mechanisms behind inverse design results [124].

### C. Deep Learning for Interpretation of Photonics Experiments

The last section of this review is dedicated to recent applications of deep learning in nano-photonics experiments.

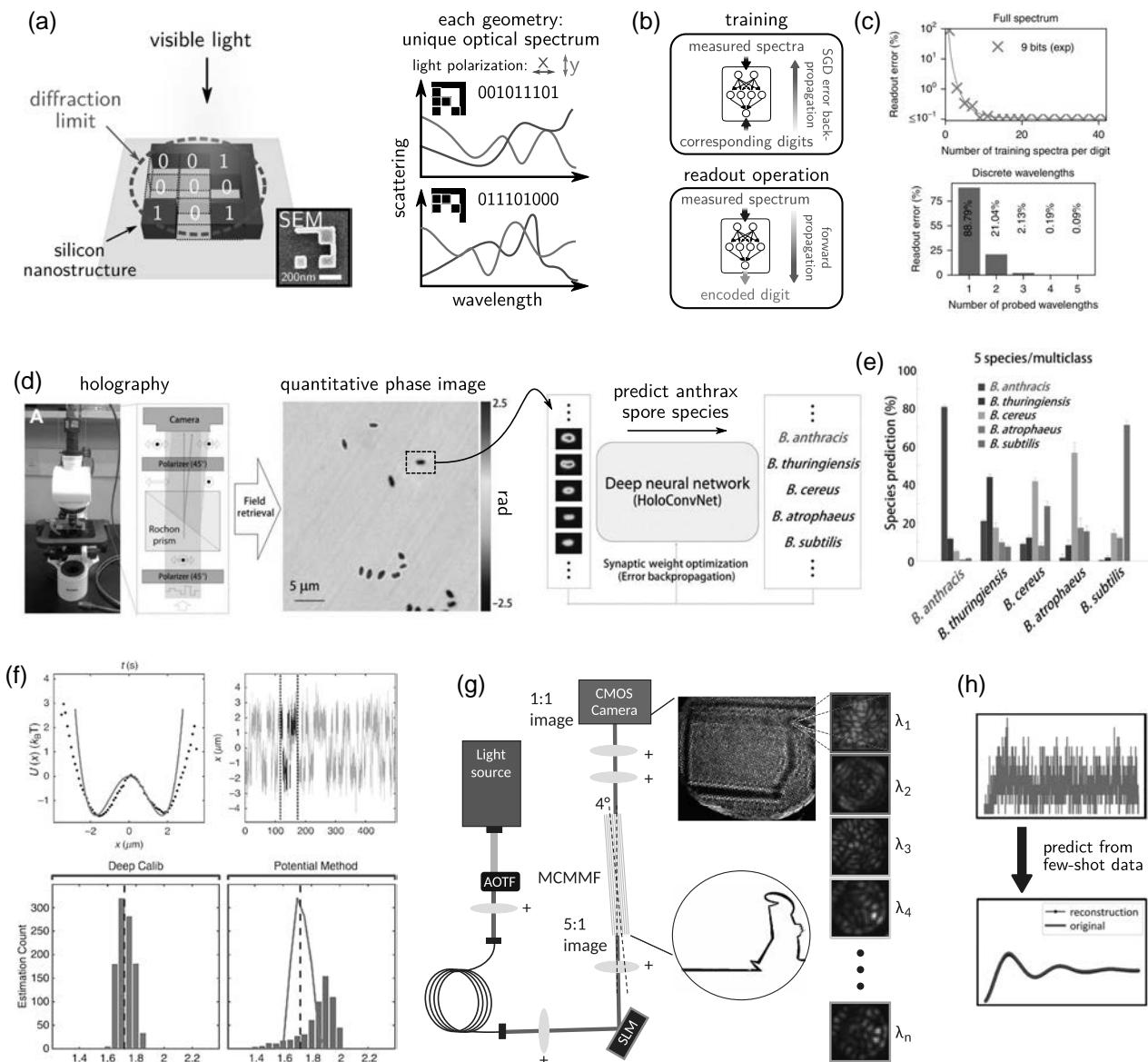
Deep learning has proven to enable unprecedented statistical evaluation of large and complicated data, which was formerly impossible with conventional methods. It has been demonstrated, for instance, that ANN models can learn from huge microscopy datasets to optically characterize 2D materials such as graphene or transition-metal dichalcogenides [125] or to automatically localize and classify nano-scale defects [126] or to track particles in 3D space using holographic microscopy [127]. Deep learning was also successfully applied for the ultra-fast analysis of single-molecule emission patterns [128] as well as for the experimental reconstruction of quantum states for quantum optics tomography [129].

By training an ANN on large amounts of experimental optical scattering spectra from complex photonic nano-structures, recently an optical information storage concept has been proposed, able to push the data density beyond the optical diffraction limit [130]. The principle is depicted in Fig. 7(a). Digital information is encoded in silicon nano-structures, which are designed such that each nano-structure encoding a specific bit-sequence possesses a unique scattering spectrum. Visible light scattering is subsequently interpreted by an artificial neural network [Fig. 7(b)]. Training on experimental data renders this readout robust against fabrication imperfections and instrumental noise. Therein, the ANN is the key ingredient to allow high readout accuracies from distorted data [Fig. 7(c)]. Deep learning can be used for various further experimental classification tasks in nano-optics. For instance, as depicted in Fig. 7(d), it has been recently demonstrated that an ANN can learn to classify different species of anthrax spores from holographic microscopy images [131]. The confusion rates in the individual classes [Fig. 7(e)] allow furthermore to assess similarities and differences between the different anthrax species. Similar recent deep-learning-based holographic image classification tasks include analysis of colloidal dispersions [132] or the real-time determination of size and refractive index of sub-wavelength small particles [133].

Deep learning is particularly strong at the interpretation of sparse, undersampled data. In a recent example, Argun *et al.* used a deep neural network for force field calibration in microscopy, by monitoring and interpreting Brownian particle motion [134]. As depicted in Fig. 7(f), complex trapping potentials (top left) can be reconstructed efficiently from few experimental samples (top right). In contrast to a conventional method (bottom right), the ANN (bottom left) reconstructs the correct potential with high accuracy also from little data [using only the dark part in the top right panel of Fig. 7(f)]. Similarly, machine learning has been used for real-time particle tracking [135–137]. Recently ANNs have also been successfully trained on simulated data to efficiently predict the optical forces in complex particle trapping situations [138]. Moreover, deep learning has been found to be very powerful in solving inverse

problems occurring in imaging experiments. In this context often sparsity assumptions are required to enable deconstruction of undersampled data, which demands computationally complex inverse solving techniques such as compressive sensing. Corresponding imaging applications include phase recovery [139,140], image reconstruction or enhancement [141–144], super-resolution microscopy [145–149], and coherent diffractive imaging [150,151]. In the context of photonics, it has been demonstrated that speckle patterns which occur after light transmission through complex media can be deconstructed very efficiently with deep learning methods [104,152–156]. While such speckles appear as if they were random patterns, they are actually the result of deterministic multiple scattering events. Therefore, a fixed correlation between input and output before and after the complex medium can be established, which is classically done by constructing a transmission matrix [157], involving complex regularization schemes, inversion procedures, or computationally expensive compressive sensing techniques [158]. While speckle-based methods allow, for instance, imaging through opaque media or the reconstruction of spectral information, the aforementioned computational burden usually prohibits real-time applications. ANN models, on the other hand, can be trained to solve the implicit inverse problem in speckle deconstruction very efficiently, which recently enabled use of complex media such as multi-mode fibers for real-time applications in imaging [104,153–155,159], spectral reconstruction [156], or both (hyper-spectral imaging) [152]. Figure 7(g) illustrates a setup for such speckle-based hyper-spectral imaging. An image is formed via an intensity spatial light modulator, spectrally shaped using an acousto-optic tunable filter, and focused on the aperture of a multi-core multi-mode fiber bundle. The fiber cores act as pixels of the image, whose individual speckle patterns encode the spectral information. Kürüm *et al.* [152] demonstrated that even under noisy conditions and in the undersampling regime, an ANN can reconstruct the spectral information of several thousand fibers with a speed of a few frames per second. In contrast, conventional compressive sensing algorithms require tens of minutes for the same task with similar reconstruction fidelity [158].

In the context of sparse data reconstruction, deep learning has recently been used in quantum optics applications for the reconstruction of statistical distributions from experiments with weak photon counts, as schematized in Fig. 7(h). For instance, Cortes *et al.* [160] demonstrated the successful reconstruction of time-dependent data from few photon events using statistical learning. In this procedure a machine learning algorithm learns to predict the statistical distribution of the data. A similar approach has been applied to assess whether a nano-diamond contains a single or several nitrogen vacancy photon emitters [161]. Another work demonstrated a machine learning model capable of differentiating between coherent and thermal light sources via a statistical analysis of the temporal distribution of a very low number of photons [162]. These learning-based statistical analysis methods are capable of outperforming conventional data fitting techniques thanks to their capacity to learn the most probable statistical distributions from the actual data. Essentially, the machine learning model learns



**Fig. 7.** Examples of ML applications in experimental data interpretation. (a)–(c) ANN used to decode information from optical information storage via a spectral scattering analysis from sub-diffraction small nano-structures. (a) Each bit sequence is encoded by a specific geometry which is designed such that it possesses a unique scattering spectrum. (b) A neural network is trained on a large amount of spectra such that it learns to decode noisy spectra of formerly not seen structures. (c) Even if only few wavelengths are probed, the readout accuracy of the network is excellent. Adapted with permission from [130], copyright (2019) Springer Nature. (d), (e) Holographic anthrax spore classification via holography microscopy. A machine learning algorithm is trained on phase images of different spore species, as depicted in (d). The neural network is capable to classify five different anthrax species with a very high accuracy. Adapted from [131], copyright (2017) the authors. (f) Microscopy force field calibration (top left, green line: trapping potential; dots: reconstructed potential). Evaluation of  $U(x)$  via ANN-based analysis of Brownian motion from undersampled statistical data (top right). Comparison of reconstruction fidelity of ANN (bottom left) and conventional method (bottom right). Ground truth is indicated by a black dashed line. Adapted from [134], copyright (2020) the authors. (g) ANN enabled real-time hyper-spectral image reconstruction from speckle patterns produced by a multi-core multi-mode fiber bundle (MCMMF). The technique exploits the wavelength dependence of the speckle patterns. Adapted from [152], copyright (2019) Optical Society of America. (h) Scheme depicting the use of machine learning for statistics reconstruction of few-shot data acquisitions. Reprinted from [160], with the permission of AIP Publishing.

to “focus” on the important regions in the data (comparable to adaptive fitting weights). Conventional data fitting algorithms on the other hand tend to attaching too much importance to “flat” areas, to the detriment of the accuracy in the relevant regions. Just as with accidentally over-specialized inverse networks, care must be taken when interpreting the ANN

reconstructions. Since data-driven approaches always bear the risk of being biased toward the training data, a neural network might, for instance, detect a learned statistical distribution even in pure noise.

Deep learning can be applied not only to data analysis but is also increasingly used to control real-time experimental

feedback systems. Recent examples touching the field of nanophotonics are mainly found in AI-stabilized microscopy. ANNs can be applied, for instance, to real-time image enhancement [163], microscopy stabilizing feedback systems [20,164], or to conduct sparse data acquisition schemes for the acceleration of scanning microscopy systems via compressive sensing [165]. ANNs have been also applied to controlling laser mode-locking stabilization systems [166–168]. So far, the direct application of ANNs to experimental hardware for nano-photonics is still scarce, but the research is in an early stage. A recent work proposed, for instance, to calibrate and control electrically reconfigurable photonic circuits by deep learning algorithms [169]. Another example is a pioneering work of Selle *et al.* [51] which proposed to use ANNs coupled to a femtosecond laser pulse shaper for real-time control of the light-matter interaction in nano-structures or molecules. We expect a very rapid development of applications in this direction in the near future; in particular, real-time critical applications such as sensing [170] will hugely benefit from the tremendous acceleration potential of ANNs.

#### 4. CONCLUSIONS AND PERSPECTIVES

In conclusion, in this mini-review we discussed the most recent developments in deep learning methods applied to nanophotonics. In the first section we focused on ANN-driven nano-photonic inverse design methods and discussed concepts to improve the design quality of inverse ANNs in comparison with conventional optimization techniques. In the second part we discussed applications of deep learning in nano-photonics “beyond inverse design,” spanning from physics-informed neural networks over ANNs for physical knowledge extraction to data interpretation and experimental applications.

We would like to emphasize that despite their latest remarkable success and their undeniable great potential, artificial neural networks are “black boxes.” It is extremely hard, mostly even impossible, to understand how a neural network generates its predictions. It has been demonstrated on many occasions that even the most sophisticated ANNs, trained on the most carefully assembled datasets, contain singular points at which their predictions diverge. Another noteworthy danger of data-driven techniques is that they bear a considerable risk to be biased with respect to their training data, such as an incident where Google’s image-tagging algorithm learned implicit racism from its training data [171]. We therefore appeal to the reader to keep in mind that, simply speaking, “what you put in is what you get out.” In consequence the ANN models are only the second most important ingredient to deep learning. The essential element is first of all the training data. Unfortunately, it is often understated and not discussed with sufficient emphasis that high-quality training data is of the utmost importance. By reviewing techniques that aim at improving the training data quality, we tried to arouse some awareness in this respect. Another important aspect in this context is the amount of training data required to train a well-performing and generalizing ANN. Unfortunately, in many problems which would be naturally suited for deep learning applications, training data is scarce or very expensive to generate. Additionally, the more general a problem for an ANN is, the more training data is

usually required for a good prediction fidelity. Last but not least, adapting an ANN model to a new problem often requires the entire training data to be generated from scratch, which might even be the case for minor modifications. These aspects can create considerable computational barriers for broad and flexible applications of ANNs.

Deep learning techniques in the context of nano-photonics have experienced a tremendous amount of attention in the past few years and research has literally exploded. ANNs have enabled manifold applications which formerly seemed strictly impossible. As discussed above, a prominent example is data-driven ultra-fast solvers for various inverse problems, for which conventional methods are computationally extremely expensive and slow. We expect that further groundbreaking applications will be developed in the near future. For instance, very promising progress has been made in the field of quantum machine learning [172], which aims at using deep learning concepts to push the capabilities and interpretability of quantum computing systems. In this context, machine learning algorithms recently have autonomously proposed designs for non-trivial quantum optics experiments [173–175]. We expect that deep learning will continue to produce exciting pioneering results. We also anticipate that deep learning techniques will become a common numerical tool, regularly employed for the daily use.

**Funding.** CALMIP Toulouse (p20010); Engineering and Physical Sciences Research Council (EP/M009122/1); Deutsche Forschungsgemeinschaft (WI 5261/1-1).

**Acknowledgment.** We thank the NVIDIA Corporation for the donation of a Quadro P6000 GPU used for this research. This work was supported by the German Research Foundation (DFG) through a research fellowship. The authors acknowledge the CALMIP computing facility. OM acknowledges support through EPSRC.

**Disclosures.** The authors declare no conflicts of interest.

#### REFERENCES

1. P. Mühlriegel, H.-J. Eisler, O. J. F. Martin, B. Hecht, and D. W. Pohl, “Resonant optical antennas,” *Science* **308**, 1607–1609 (2005).
2. A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk'yanchuk, “Optically resonant dielectric nanostructures,” *Science* **354**, aag2472 (2016).
3. C. Girard, “Near fields in nanostructures,” *Rep. Prog. Phys.* **68**, 1883–1933 (2005).
4. M. Kauranen and A. V. Zayats, “Nonlinear plasmonics,” *Nat. Photonics* **6**, 737–748 (2012).
5. G. C. des Francs, J. Barthes, A. Bouhelier, J. C. Weeber, A. Dereux, A. Cuche, and C. Girard, “Plasmonic Purcell factor and coupling efficiency to surface plasmons. Implications for addressing and controlling optical nanosources,” *J. Opt.* **18**, 094005 (2016).
6. J. Wang, F. Sciarrino, A. Laing, and M. G. Thompson, “Integrated photonic quantum technologies,” *Nat. Photonics* **14**, 273–284 (2020).
7. J. B. Pendry, “Negative refraction makes a perfect lens,” *Phys. Rev. Lett.* **85**, 3966–3969 (2000).
8. P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, and R. Devlin, “Recent advances in planar optics: from plasmonic to dielectric metasurfaces,” *Optica* **4**, 139–152 (2017).

9. N. M. Estakhri, B. Edwards, and N. Engheta, "Inverse-designed metastructures that solve equations," *Science* **363**, 1333–1338 (2019).
10. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, and I. A. Walmsley, "Optimal design for universal multiport interferometers," *Optica* **3**, 1460–1465 (2016).
11. F. Zangeneh-Nejad, D. L. Sounas, A. Alù, and R. Fleury, "Analogue computing with metamaterials," *Nat. Rev. Mater.* **6**, 207–225 (2021).
12. E. A. Ash and G. Nicholls, "Super-resolution aperture scanning microscope," *Nature* **237**, 510–512 (1972).
13. D. W. Pohl, W. Denk, and M. Lanz, "Optical stethoscopy: image recording with resolution  $\lambda/20$ ," *Appl. Phys. Lett.* **44**, 651–653 (1984).
14. E. Betzig, A. Harootunian, A. Lewis, and M. Isaacson, "Near-field diffraction by a slit: implications for superresolution microscopy," *Appl. Opt.* **25**, 1890–1900 (1986).
15. B. Gallinet, J. Butet, and O. J. F. Martin, "Numerical methods for nanophotonics: standard problems and future challenges," *Laser Photonics Rev.* **9**, 577–603 (2015).
16. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
17. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
18. S. Chan and E. L. Siegel, "Will machine learning end the viability of radiology as a thriving medical specialty?" *Br. J. Radiol.* **92**, 20180416 (2018).
19. A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.* **29**, 102–127 (2019).
20. D. A. Cirovic, "Feed-forward artificial neural networks: applications to spectroscopy," *TRAC Trends Anal. Chem.* **16**, 148–155 (1997).
21. T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. M. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever, and D. Amodei, "Language models are few-shot learners," in *Proceedings of Advances in Neural Information Processing System* (2020), pp. 1877–1901.
22. C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence* (2016), pp. 4278–4284.
23. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, "All-optical machine learning using diffractive deep neural networks," *Science* **361**, 1004–1008 (2018).
24. T. W. Hughes, I. A. D. Williamson, M. Minkov, and S. Fan, "Wave physics as an analog recurrent neural network," *Sci. Adv.* **5**, eaay6946 (2019).
25. D. Mengu, Y. Rivenson, and A. Ozcan, "Scale-, shift- and rotation-invariant diffractive optical networks," *ACS Photon.* **8**, 324–334 (2021).
26. R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.* **2**, 1007–1023 (2020).
27. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2020).
28. S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **9**, 1041–1057 (2020).
29. J. Jiang, M. Chen, and J. A. Fan, "Deep neural networks for the evaluation and design of photonic devices," *arXiv:2007.00084* (2020).
30. L. Huang, L. Xu, and A. E. Miroshnichenko, "Deep learning enabled nanophotonics," in *Advances in Deep Learning* (InTech, 2020).
31. M. M. R. Elsayy, S. Lanteri, R. Duvigneau, J. A. Fan, and P. Genevet, "Numerical optimization methods for metasurfaces," *Laser Photonics Rev.* **14**, 1900445 (2020).
32. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
33. G. M. Sacha and P. Varona, "Artificial intelligence in nanotechnology," *Nanotechnology* **24**, 452002 (2013).
34. K. Yao, R. Unni, and Y. Zheng, "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**, 339–366 (2019).
35. J. Zhou, B. Huang, Z. Yan, and J.-C. G. Bünzli, "Emerging role of machine learning in light-matter interaction," *Light Sci. Appl.* **8**, 1 (2019).
36. D. Piccinotti, K. F. MacDonald, S. Gregory, I. Youngs, and N. I. Zheludev, "Artificial intelligence for photonics and photonic materials," *Rep. Prog. Phys.* **84**, 012401 (2020).
37. J. Moughames, X. Porte, M. Thiel, G. Ulliac, L. Larger, M. Jacquot, M. Kadic, and D. Brunner, "Three-dimensional waveguide interconnects for scalable integration of photonic neural networks," *Optica* **7**, 640–646 (2020).
38. X. Porte, A. Skalli, N. Haghghi, S. Reitzenstein, J. A. Lott, and D. Brunner, "A complete, parallel and autonomous photonic neural network in a semiconductor multimode laser," *arXiv:2012.11153* (2020).
39. L.-J. Black, Y. Wang, C. H. de Groot, A. Arbouet, and O. L. Muskens, "Optimal polarization conversion in coupled dimer plasmonic nanoantennas for metasurfaces," *ACS Nano* **8**, 6390–6399 (2014).
40. M. Celebrano, X. Wu, M. Baselli, S. Großmann, P. Biagioni, A. Locatelli, C. De Angelis, G. Cerullo, R. Osellame, B. Hecht, L. Duò, F. Cicaccia, and M. Finazzi, "Mode matching in multiresonant plasmonic nanoantennas for enhanced second harmonic generation," *Nat. Nanotechnol.* **10**, 412–417 (2015).
41. J. S. Jensen and O. Sigmund, "Topology optimization for nanophotonics," *Laser Photonics Rev.* **5**, 308–321 (2011).
42. S. D. Campbell, D. Sell, R. P. Jenkins, E. B. Whiting, J. A. Fan, and D. H. Werner, "Review of numerical optimization techniques for meta-device design [Invited]," *Opt. Mater. Express* **9**, 1842–1863 (2019).
43. F. Meng, X. Huang, and B. Jia, "Bi-directional evolutionary optimization for photonic band gap structures," *J. Comput. Phys.* **302**, 393–404 (2015).
44. S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.* **79**, 12–49 (1988).
45. T. Feichtner, O. Selig, M. Kiunke, and B. Hecht, "Evolutionary optimization of optical antennas," *Phys. Rev. Lett.* **109**, 127701 (2012).
46. P. R. Wiecha, P. R. Wiecha, C. Majorel, C. Girard, A. Cuche, V. Paillard, O. L. Muskens, and A. Arbouet, "Design of plasmonic directional antennas via evolutionary optimization," *Opt. Express* **27**, 29069–29081 (2019).
47. P. R. Wiecha, A. Arbouet, C. Girard, A. Lecestre, G. Larrieu, and V. Paillard, "Evolutionary multi-objective optimization of colour pixels based on dielectric nanoantennas," *Nat. Nanotechnol.* **12**, 163–169 (2017).
48. D. Z. Zhu, E. B. Whiting, S. D. Campbell, D. B. Burckel, and D. H. Werner, "Optimal high efficiency 3D plasmonic metasurface elements revealed by lazy ants," *ACS Photonics* **6**, 2741–2748 (2019).
49. Y. Augenstein and C. Rockstuhl, "Inverse design of nanophotonic devices with structural integrity," *ACS Photonics* **7**, 2190–2196 (2020).
50. R. Selle, G. Vogt, T. Brixner, G. Gerber, R. Metzler, and W. Kinzel, "Modeling of light-matter interactions with neural networks," *Phys. Rev. A* **76**, 023810 (2007).
51. R. Selle, T. Brixner, T. Bayer, M. Wollenhaupt, and T. Baumert, "Modelling of ultrafast coherent strong-field dynamics in potassium with neural networks," *J. Phys. B* **41**, 074019 (2008).
52. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
53. S. An, C. Fowler, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, C. Rivero-Baleine, K. A. Richardson, T. Gu, J. Hu, and H. Zhang, "A deep learning approach for objective-driven all-dielectric metasurface design," *ACS Photonics* **6**, 3196–3207 (2019).
54. M. V. Zhelyeznyakov, S. L. Brunton, and A. Majumdar, "Deep learning to accelerate Maxwell's equations for inverse design of dielectric metasurfaces," *arXiv:2008.10632* (2020).

55. Y. Li, Y. Wang, S. Qi, Q. Ren, L. Kang, S. D. Campbell, P. L. Werner, and D. H. Werner, "Predicting scattering from complex nanostructures via deep learning," *IEEE Access* **8**, 139983 (2020).
56. S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core-shell nanoparticles," *ACS Appl. Mater. Interfaces* **11**, 24264–24268 (2019).
57. A.-P. Blanchard-Dionne and O. J. F. Martin, "Teaching optics to a machine learning network," *Opt. Lett.* **45**, 2922–2925 (2020).
58. P. R. Wiecha and O. L. Muskens, "Deep learning meets nanophotonics: a generalized accurate predictor for near fields and far fields of arbitrary 3D nanostructures," *Nano Lett.* **20**, 329–338 (2020).
59. Y. Zhu, N. Zabaras, P.-S. Koutsourelakis, and P. Perdikaris, "Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data," *J. Comput. Phys.* **394**, 56–81 (2019).
60. T. Chugh, C. Sun, H. Wang, and Y. Jin, "Surrogate-assisted evolutionary optimization of large problems," in *High-Performance Simulation-Based Optimization*, T. Bartz-Beielstein, B. Filipič, P. Korošec, and E.-G. Talbi, eds. (Springer, 2020), pp. 165–187.
61. S. D. Campbell, D. Z. Zhu, E. B. Whiting, J. Nagar, D. H. Werner, and P. L. Werner, "Advanced multi-objective and surrogate-assisted optimization of topologically diverse metasurface architectures," *Proc. SPIE* **10719**, 107190U (2018).
62. V. Kalt, A. K. González-Alcalde, S. Es-Saidi, R. Salas-Montiel, S. Blaize, and D. Macías, "Metamodeling of high-contrast-index gratings for color reproduction," *J. Opt. Soc. Am. A* **36**, 79–88 (2019).
63. A. K. González-Alcalde, R. Salas-Montiel, V. Kalt, S. Blaize, and D. Macías, "Engineering colors in all-dielectric metasurfaces: metamodelling approach," *Opt. Lett.* **45**, 89–92 (2020).
64. R. Pestourie, Y. Mroueh, T. V. Nguyen, P. Das, and S. G. Johnson, "Active learning of deep surrogates for PDEs: application to metasurface design," *arXiv:2008.12649* (2020).
65. R. S. Hegde, "Photonics inverse design: pairing deep neural networks with evolutionary algorithms," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700908 (2020).
66. Z. A. Kudyshev, A. V. Kildishev, V. M. Shalaev, and A. Boltasseva, "Machine learning assisted global optimization of photonic devices," *Nanophotonics* **10**, 371–383 (2020).
67. J. Su, D. V. Vargas, and S. Kouichi, "One pixel attack for fooling deep neural networks," *IEEE Trans. Evol. Comput.* **23**, 828–841 (2019).
68. J. Jiang, D. Sell, S. Hoyer, J. Hickey, J. Yang, and J. A. Fan, "Freeform diffractive metagrating design based on generative adversarial networks," *ACS Nano* **13**, 8872–8878 (2019).
69. R. Trivedi, L. Su, J. Lu, M. F. Schubert, and J. Vuckovic, "Data-driven acceleration of photonic simulations," *Sci. Rep.* **9**, 19728 (2019).
70. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics* **5**, 1365–1369 (2018).
71. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
72. L. Gao, X. Li, D. Liu, L. Wang, and Z. Yu, "A bidirectional deep neural network for accurate silicon color design," *Adv. Mater.* **31**, 1905467 (2019).
73. S. An, B. Zheng, H. Tang, M. Y. Shalaginov, L. Zhou, H. Li, T. Gu, J. Hu, C. Fowler, and H. Zhang, "Multifunctional metasurface design with a generative adversarial network," *arXiv:1908.04851* (2020).
74. Z. Liu, D. Zhu, S. P. Rodrigues, K.-T. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
75. S. So and J. Rho, "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**, 1255–1261 (2019).
76. A. Mall, A. Patil, D. Tamboli, A. Sethi, and A. Kumar, "Fast design of plasmonic metasurfaces enabled by deep learning," *J. Phys. D* **53**, 49LT01 (2020).
77. Z. Liu, L. Raju, D. Zhu, and W. Cai, "A hybrid strategy for the discovery and design of photonic structures," *IEEE J. Emerging Sel. Top. Circuits Syst.* **10**, 126–135 (2020).
78. W. Ma, F. Cheng, Y. Xu, Q. Wen, and Y. Liu, "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Adv. Mater.* **31**, 1901111 (2019).
79. X. Shi, T. Qiu, J. Wang, X. Zhao, and S. Qu, "Metasurface inverse design using machine learning approaches," *J. Phys. D* **53**, 275105 (2020).
80. W. Ma and Y. Liu, "A data-efficient self-supervised deep learning model for design and characterization of nanophotonic structures," *Sci. China Phys. Mech. Astron.* **63**, 284212 (2020).
81. D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *Found. Trends Mach. Learn.* **12**, 307–392 (2019).
82. T. Badloe, I. Kim, and J. Rho, "Biomimetic ultra-broadband perfect absorbers optimised with reinforcement learning," *Phys. Chem. Chem. Phys.* **22**, 2337–2342 (2020).
83. H. Wang, Z. Zheng, C. Ji, and L. J. Guo, "Automated multi-layer optical design via deep reinforcement learning," *Mach. Learn. Sci. Technol.* (2020).
84. H. Wang, Z. Zheng, C. Ji, and L. J. Guo, "Automated multi-layer optical design via deep reinforcement learning," *Mach. Learn. Sci. Technol.* **2**, 025013 (2021).
85. E. Ashalley, K. Acheampong, L. V. Besteiro, L. V. Besteiro, P. Yu, A. Neogi, A. O. Govorov, A. O. Govorov, and Z. M. Wang, "Multitask deep-learning-based design of chiral plasmonic metamaterials," *Photon. Res.* **8**, 1213–1225 (2020).
86. J. Trisno, H. Wang, H. T. Wang, R. J. H. Ng, S. D. Rezaei, and J. K. W. Yang, "Applying machine learning to the optics of dielectric nanoblobs," *Adv. Photonics Res.* **1**, 2000068 (2020).
87. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
88. A. Sheverdin, F. Monticone, and C. Valagiannopoulos, "Photonic inverse design with neural networks: the case of invisibility in the visible," *Phys. Rev. Appl.* **14**, 024054 (2020).
89. A.-P. Blanchard-Dionne and O. J. F. Martin, "Successive training of a generative adversarial network for the design of an optical cloak," *OSA Contin.* **4**, 87–95 (2021).
90. Y. Chen, L. Lu, G. E. Karniadakis, and L. D. Negro, "Physics-informed neural networks for inverse problems in nano-optics and metamaterials," *Opt Express* **28**, 11618–11633 (2020).
91. I. Sajedian, H. Lee, and J. Rho, "Double-deep Q-learning to increase the efficiency of metasurface holograms," *Sci. Rep.* **9**, 10899 (2019).
92. A. D. Phan, C. V. Nguyen, P. T. Linh, T. V. Huynh, V. D. Lam, A.-T. Le, and K. Wakabayashi, "Deep learning for the inverse design of mid-infrared graphene plasmons," *Crystals* **10**, 125 (2020).
93. C. Yeung, J.-M. Tsai, B. King, B. Pham, J. Liang, D. Ho, M. W. Knight, and A. P. Raman, "Designing multiplexed supercell metasurfaces with tandem neural networks," *Nanophotonics* **10**, 1133–1143 (2021).
94. T. Asano and S. Noda, "Iterative optimization of photonic crystal nanocavity designs by using deep neural networks," *Nanophotonics* **8**, 2243–2256 (2019).
95. F. Wen, J. Jiang, and J. A. Fan, "Robust freeform metasurface design based on progressively growing generative networks," *ACS Photonics* **7**, 2098–2104 (2020).
96. S. Wang, K. Fan, N. Luo, Y. Cao, F. Wu, C. Zhang, K. A. Heller, and L. You, "Massive computational acceleration by using neural networks to emulate mechanism-based biological models," *Nat. Commun.* **10**, 4354 (2019).
97. J. Jiang and J. A. Fan, "Multiobjective and categorical global optimization of photonic structures based on ResNet generative neural networks," *Nanophotonics* **10**, 361–369 (2020).
98. J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.* **19**, 5366–5372 (2019).
99. K. Deb, *Multi-Objective Optimization Using Evolutionary Algorithms* (Wiley, 2001), Vol. **16**.
100. R. Unni, K. Yao, and Y. Zheng, "Deep convolutional mixture density network for inverse design of layered photonic structures," *ACS Photonics* **7**, 2703–2712 (2020).

101. B. Hu, B. Wu, D. Tan, J. Xu, J. Xu, Y. Chen, and Y. Chen, "Robust inverse-design of scattering spectrum in core-shell structure using modified denoising autoencoder neural network," *Opt. Express* **27**, 36276–36285 (2019).
102. Z. Fang and J. Zhan, "Deep physical informed neural networks for metamaterial design," *IEEE Access* **8**, 24506–24513 (2020).
103. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," *arXiv:1505.04597* (2015).
104. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).
105. H. Kabir, Y. Wang, M. Yu, and Q.-J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microwave Theory Tech.* **56**, 867–879 (2008).
106. C. Zhang, J. Jin, W. Na, Q.-J. Zhang, and M. Yu, "Multivalued neural network inverse modeling and applications to microwave filters," *IEEE Trans. Microwave Theory Tech.* **66**, 3781–3797 (2018).
107. Y.-T. Luo, P.-Q. Li, D.-T. Li, Y.-G. Peng, Z.-G. Geng, S.-H. Xie, Y. Li, A. Alù, J. Zhu, and X.-F. Zhu, "Probability-density-based deep learning paradigm for the fuzzy design of functional metasurfaces," *Research* **2020**, 8757403 (2020).
108. J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Advances in Neural Information Processing Systems*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds. (Curran Associates, 2012), Vol. **25**, pp. 341–349.
109. H. Gao, L. Sun, and J.-X. Wang, "PhyGeoNet: physics-informed geometry-adaptive convolutional neural networks for solving parameterized steady-state PDEs on irregular domain," *J. Comput. Phys.* **428**, 110079 (2020).
110. Z. Liu, Z. Liu, Z. Zhu, and W. Cai, "Topological encoding method for data-driven photonics inverse design," *Opt. Express* **28**, 4825–4835 (2020).
111. C. C. Nadell, B. Huang, J. M. Malof, and W. J. Padilla, "Deep learning for accelerated all-dielectric metasurface design," *Opt. Express* **27**, 27523–27535 (2019).
112. Y. Qu, L. Jing, Y. Shen, M. Qiu, and M. Soljačić, "Migrating knowledge between physical scenarios based on artificial neural networks," *ACS Photonics* **6**, 1168–1174 (2019).
113. M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, "Machine learning analysis of extreme events in optical fibre modulation instability," *Nat. Commun.* **9**, 4923 (2018).
114. M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: a deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.* **378**, 686–707 (2019).
115. M. Raissi, A. Yazdani, and G. E. Karniadakis, "Hidden fluid mechanics: learning velocity and pressure fields from flow visualizations," *Science* **367**, 1026–1030 (2020).
116. B. Moseley, A. Markham, and T. Nissen-Meyer, "Solving the wave equation with physics-informed deep learning," *arXiv:2006.11894* (2020).
117. Y. Kiarashinejad, S. Abdollahramezani, M. Zandehshahvar, O. Hemmatyar, and A. Adibi, "Deep learning reveals underlying physics of light-matter interactions in nanophotonic devices," *Adv. Theor. Simul.* **2**, 1900088 (2019).
118. Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, "Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures," *arXiv:1902.03865* (2019).
119. C. Yeung, J.-M. Tsai, B. King, Y. Kawagoe, D. Ho, M. Knight, and A. P. Raman, "Elucidating the behavior of nanophotonic structures through explainable machine learning algorithms," *ACS Photonics* **7**, 2309–2318 (2020).
120. R. Iten, T. Metger, H. Wilming, L. del Rio, and R. Renner, "Discovering physical concepts with neural networks," *Phys. Rev. Lett.* **124**, 010508 (2020).
121. Y. Kiarashinejad, M. Zandehshahvar, S. Abdollahramezani, O. Hemmatyar, R. Pourabolghasem, and A. Adibi, "Knowledge discovery in nanophotonics using geometric deep learning," *Adv. Intell. Syst.* **2**, 1900132 (2020).
122. S. An, B. Zheng, M. Y. Shalaginov, H. Tang, H. Li, L. Zhou, J. Ding, A. M. Agarwal, A. M. Agarwal, C. Rivero-Baleine, M. Kang, K. A. Richardson, T. Gu, J. Hu, C. Fowler, C. Fowler, H. Zhang, and H. Zhang, "Deep learning modeling approach for metasurfaces with high degrees of freedom," *Opt. Express* **28**, 31932–31942 (2020).
123. C. Yeung, J.-M. Tsai, Y. Kawagoe, B. King, D. Ho, and A. P. Raman, "Elucidating the design and behavior of nanophotonic structures through explainable convolutional neural networks," *arXiv:2003.06075* (2020).
124. M. Elzouka, C. Yang, A. Albert, S. Lubner, and R. S. Prasher, "Interpretable inverse design of particle spectral emissivity using machine learning," *arXiv:2002.04223* (2020).
125. B. Han, Y. Lin, Y. Yang, N. Mao, W. Li, H. Wang, V. Fatemi, L. Zhou, J. I.-J. Wang, Q. Ma, Y. Cao, D. Rodan-Legrain, Y.-Q. Bie, E. Navarro-Moratalla, D. Klein, D. MacNeill, S. Wu, W. S. Leong, H. Kitadai, X. Ling, P. Jarillo-Herrero, T. Palacios, J. Yin, and J. Kong, "Deep learning enabled fast optical characterization of two-dimensional materials," *arXiv:1906.11220* (2019).
126. M. Ziatdinov, O. Dyck, A. Maksov, X. Li, X. Sang, K. Xiao, R. R. Unocic, R. Vasudevan, S. Jesse, and S. V. Kalinin, "Deep learning of atomically resolved scanning transmission electron microscopy images: chemical identification and tracking local transformations," *ACS Nano* **11**, 12742–12752 (2017).
127. S. Shao, S. Shao, K. Mallery, K. Mallery, S. S. Kumar, S. S. Kumar, J. Hong, and J. Hong, "Machine learning holography for 3D particle field imaging," *Opt. Express* **28**, 2987–2999 (2020).
128. P. Zhang, S. Liu, A. Chaurasia, D. Ma, M. J. Mlodzianoski, E. Culurciello, and F. Huang, "Analyzing complex single-molecule emission patterns with deep learning," *Nat. Methods* **15**, 913–916 (2018).
129. A. M. Palmieri, E. Kovlakov, F. Bianchi, D. Yudin, S. Straupe, J. D. Biamonte, and S. Kulik, "Experimental neural network enhanced quantum tomography," *npj Quantum Inf.* **6**, 20 (2020).
130. P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nat. Nanotechnol.* **14**, 237–244 (2019).
131. Y. Jo, S. Park, J. Jung, J. Yoon, H. Joo, M.-H. Kim, S.-J. Kang, M. C. Choi, S. Y. Lee, and Y. Park, "Holographic deep learning for rapid optical screening of anthrax spores," *Sci. Adv.* **3**, e1700606 (2017).
132. A. Yevick, M. Hannel, and D. G. Grier, "Machine-learning approach to holographic particle characterization," *Opt. Express* **22**, 26884–26890 (2014).
133. B. Midtvedt, E. Olsén, F. Eklund, F. Höök, C. B. Adiels, G. Volpe, and D. Midtvedt, "Holographic characterisation of subwavelength particles enhanced by deep learning," *arXiv:2006.11154* (2020).
134. A. Argun, T. Thalheim, S. Bo, F. Cichos, and G. Volpe, "Enhanced force-field calibration via machine learning," *Appl. Phys. Rev.* **7**, 041404 (2020).
135. M. D. Hannel, A. Abdulali, M. O'Brien, and D. G. Grier, "Machine-learning techniques for fast and accurate feature localization in holograms of colloidal particles," *Opt. Express* **26**, 15221–15231 (2018).
136. J. M. Newby, A. M. Schaefer, P. T. Lee, M. G. Forest, and S. K. Lai, "Convolutional neural networks automate detection for tracking of submicron-scale particles in 2D and 3D," *Proc. Natl. Acad. Sci. USA* **115**, 9026–9031 (2018).
137. S. Helgadottir, A. Argun, and G. Volpe, "Digital video microscopy enhanced by deep learning," *Optica* **6**, 506–513 (2019).
138. I. C. D. Lenton, G. Volpe, A. B. Stilgoe, T. A. Nieminen, and H. Rubinsztein-Dunlop, "Machine learning reveals complex behaviours in optically trapped particles," *Mach. Learn. Sci. Technol.* **1**, 045009 (2020).
139. Y. Rivenson, Y. Zhang, H. Günaydn, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**, 17141 (2018).
140. Y. Nishizaki, R. Horisaki, K. Kitaguchi, M. Saito, and J. Tanida, "Analysis of non-iterative phase retrieval based on machine learning," *Opt. Rev.* **27**, 136–141 (2020).
141. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).

142. G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE J. Sel. Areas Inform. Theor.* **1**, 39–56 (2020).
143. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
144. Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, "Deep learning microscopy," *Optica* **4**, 1437–1443 (2017).
145. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: super-resolution single-molecule microscopy by deep learning," *Optica* **5**, 458–464 (2018).
146. W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," *Nat. Biotechnol.* **36**, 460–468 (2018).
147. E. Nehme, D. Freedman, R. Gordon, B. Ferdinand, L. E. Weiss, O. Alalouf, T. Naor, R. Orange, T. Michaeli, and Y. Shechtman, "DeepSTORM3D: dense 3D localization microscopy and PSF design by deep learning," *Nat. Methods* **17**, 734–740 (2020).
148. T. Pu, J.-Y. Ou, V. Savinov, G. Yuan, N. Papasimakis, and N. Zheludev, "Unlabeled far-field deeply subwavelength topological microscopy (DSTM)," *Adv. Sci.* **8**, 2002886 (2020).
149. T. Pu, J. Y. Ou, N. Papasimakis, and N. I. Zheludev, "Label-free deeply subwavelength optical microscopy," *Appl. Phys. Lett.* **116**, 131105 (2020).
150. D. Bouchet, J. Seifert, and A. P. Mosk, "Optimizing illumination for precise multi-parameter estimations in coherent diffractive imaging," *Opt. Lett.* **46**, 254–257 (2021).
151. A. Ghosh, D. J. Roth, L. H. Nicholls, W. P. Wardley, A. V. Zayats, and V. A. Podolskiy, "Machine learning—based diffractive imaging with subwavelength resolution," arXiv:2005.03595 (2020).
152. U. Kürüm, P. R. Wiecha, R. French, and O. L. Muskens, "Deep learning enabled real time speckle recognition and hyperspectral imaging using a multimode fiber array," *Opt. Express* **27**, 20965–20979 (2019).
153. R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," *Opt. Express* **24**, 13738–13743 (2016).
154. L. Yunzhe, X. Yujia, and T. Lei, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–11819 (2018).
155. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," *Light Sci. Appl.* **7**, 69 (2018).
156. G. D. Bruce, L. O'Donnell, M. Chen, M. Facchin, and K. Dholakia, "Femtometer-resolved simultaneous measurement of multiple laser wavelengths in a speckle wavemeter," *Opt. Lett.* **45**, 1926–1929 (2020).
157. S. Popoff, G. Lerosey, M. Fink, A. C. Boccara, and S. Gigan, "Image transmission through an opaque material," *Nat. Commun.* **1**, 81 (2010).
158. R. French, S. Gigan, and O. L. Muskens, "Snapshot fiber spectral imaging using speckle correlations and compressive sensing," *Opt. Express* **26**, 32302–32316 (2018).
159. H. Pinkard, Z. Phillips, A. Babakhani, D. A. Fletcher, and L. Waller, "Deep learning for single-shot autofocus microscopy," *Optica* **6**, 794–797 (2019).
160. C. L. Cortes, S. Adhikari, X. Ma, and S. K. Gray, "Accelerating quantum optics experiments with statistical learning," *Appl. Phys. Lett.* **116**, 184003 (2020).
161. Z. A. Kudyshev, S. I. Bogdanov, T. Isacsson, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, "Rapid classification of quantum sources enabled by machine learning," *Adv. Quantum Technol.* **3**, 2000067 (2020).
162. C. You, M. A. Quiroz-Juárez, A. Lambert, N. Bhusal, C. Dong, A. Perez-Leija, A. Javaid, R. de. J. León-Montiel, and O. S. Magaña-Loaiza, "Identification of light sources using machine learning," *Appl. Phys. Rev.* **7**, 021404 (2020).
163. Y. Rivenson, H. Ceylan Koydemir, H. Wang, Z. Wei, Z. Ren, H. Günaydin, Y. Zhang, Z. Göröcs, K. Liang, D. Tseng, and A. Ozcan, "Deep learning enhanced mobile-phone microscopy," *ACS Photonics* **5**, 2354–2364 (2018).
164. X. Li, J. Dong, B. Li, Y. Zhang, Y. Zhang, A. Veeraraghavan, and X. Ji, "Fast confocal microscopy imaging based on deep learning," in *IEEE International Conference on Computational Photography (ICCP)* (2020), pp. 1–12.
165. J. M. Ede and R. Beanland, "Partial scanning transmission electron microscopy with deep learning," *Sci. Rep.* **10**, 8332 (2020).
166. S. L. Brunton, X. Fu, and J. N. Kutz, "Self-tuning fiber lasers," *IEEE J. Sel. Top. Quantum Electron.* **20**, 464–471 (2014).
167. J. N. Kutz and S. L. Brunton, "Intelligent systems for stabilizing mode-locked lasers and frequency combs: machine learning and equation-free control paradigms for self-tuning optics," *Nanophotonics* **4**, 459–471 (2015).
168. T. Baumeister, S. L. Brunton, and J. N. Kutz, "Deep learning and model predictive control for self-tuning mode-locked lasers," *J. Opt. Soc. Am. B* **35**, 617–626 (2018).
169. A. Youssry, R. J. Chapman, A. Peruzzo, C. Ferrie, and M. Tomamichel, "Modeling and control of a reconfigurable photonic circuit using deep learning," *Quantum Sci. Technol.* **5**, 025001 (2020).
170. B. Wang, J. C. Cancilla, J. S. Torrecilla, and H. Haick, "Artificial sensing intelligence with silicon nanowires for ultraselective detection in the gas phase," *Nano Lett.* **14**, 933–938 (2014).
171. "Google says sorry for racist auto-tag in photo app," <https://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app> (2015).
172. M. Schuld, I. Sinayskiy, and F. Petruccione, "An introduction to quantum machine learning," *Contemp. Phys.* **56**, 172–185 (2015).
173. M. Krenn, M. Malik, R. Fickler, R. Lapkiewicz, and A. Zeilinger, "Automated search for new quantum experiments," *Phys. Rev. Lett.* **116**, 090405 (2016).
174. A. A. Melnikov, H. P. Nautrup, M. Krenn, V. Dunjko, M. Tiersch, A. Zeilinger, and H. J. Briegel, "Active learning machine learns to create new quantum experiments," *Proc. Natl. Acad. Sci. USA* **115**, 1221–1226 (2018).
175. M. Krenn, M. Erhard, and A. Zeilinger, "Computer-inspired quantum experiments," *Nat. Rev. Phys.* **2**, 649–661 (2020).

# PHOTONICS Research

## All-optical neuromorphic binary convolution with a spiking VCSEL neuron for image gradient magnitudes

YAHUI ZHANG,<sup>1,2</sup> JOSHUA ROBERTSON,<sup>1</sup> SHUIYING XIANG,<sup>2,\*</sup> MATEJ HEJDA,<sup>1</sup> JULIÁN BUENO,<sup>1</sup> AND ANTONIO HURTADO<sup>1,3</sup>

<sup>1</sup>Institute of Photonics, SUPA Department of Physics, University of Strathclyde, Glasgow G1 1RD, UK

<sup>2</sup>State Key Laboratory of Integrated Service Networks, Xidian University, Xi'an 710071, China

<sup>3</sup>e-mail: antonio.hurtado@strath.ac.uk

\*Corresponding author: jxxsy@126.com

Received 9 October 2020; revised 2 February 2021; accepted 18 February 2021; posted 19 February 2021 (Doc. ID 412141); published 14 April 2021

All-optical binary convolution with a photonic spiking vertical-cavity surface-emitting laser (VCSEL) neuron is proposed and demonstrated experimentally for the first time, to the best of our knowledge. Optical inputs, extracted from digital images and temporally encoded using rectangular pulses, are injected in the VCSEL neuron, which delivers the convolution result in the number of fast (<100 ps long) spikes fired. Experimental and numerical results show that binary convolution is achieved successfully with a single spiking VCSEL neuron and that all-optical binary convolution can be used to calculate image gradient magnitudes to detect edge features and separate vertical and horizontal components in source images. We also show that this all-optical spiking binary convolution system is robust to noise and can operate with high-resolution images. Additionally, the proposed system offers important advantages such as ultrafast speed, high-energy efficiency, and simple hardware implementation, highlighting the potentials of spiking photonic VCSEL neurons for high-speed neuromorphic image processing systems and future photonic spiking convolutional neural networks. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.412141>

### 1. INTRODUCTION

Convolutional neural networks (CNNs) have seen tremendous success in many applications, such as speech and image recognition [1,2], computer vision [3], and document analysis [4]. However, CNN-based systems are computationally expensive due to their complicated architectures and the large number of parameters they rely on. CNNs therefore typically require the implementation of multicore central processing units and graphics processing units to compensate for the rather high computational expense [5,6]. This makes CNN architectures often unsuitable for smaller devices like phones and smart cameras, where power and speed have strict limitations. To address these drawbacks, the optimization and discovery of new high-speed and low power consumption platforms for CNNs are urgently required. For the optimization of CNNs, binary CNNs, which are simple, efficient, and accurate approximations of complete CNNs, can be introduced [7–9]. In binary CNNs, the weights given to the inputs of each convolutional layer are approximated with binary values [7]. Therefore, binary CNNs boast 58× faster convolutional operations and 32× less memory requirements than traditional CNNs [7].

Several optimized binary versions of CNNs have been proposed for training processes and image classification tasks [7,10,11]. However, beyond the optimization of CNNs, a new platform offering high speed and low power consumption remains highly desirable.

Photonics is considered a highly promising candidate for future neural network implementations given the unique advantages it provides such as high speed, wide bandwidth, and low power consumption [12–21]. Photonics-based CNNs have therefore been proposed in order to increase the speed of convolutional operations [18–21]. A photonic CNN accelerator was proposed based on silicon photonic micro-ring weighting banks [18]. The full system design offers more than three orders of magnitude improvement in execution time, and its optical core potentially offers more than five orders of magnitude improvement compared to state-of-the-art electronic counterparts [18]. Xu *et al.* also proposed high-accuracy optical convolution unit architecture based on acousto-optical modulator arrays, where the optical convolution unit was shown to perform well on inferences of typical CNN tasks [20]. However, the size of the system is based on the size of the kernel utilized in these emerging works on photonic CNNs.

In this work, we propose an all-optical binary convolution system using a single vertical-cavity surface-emitting laser (VCSEL) operating as a spiking optical neuron, hence, dramatically reducing hardware requirements. In our approach, temporal encoding is used instead of spatial encoding, thus crucially helping to reduce (optical) hardware complexity. In our all-optical binary convolution technique, results are represented by the number of fast ( $<100$  ps long) spiking responses delivered by the optical spiking VCSEL neuron. This has unique advantages in terms of robustness to noise and high precision. Additionally, VCSELs have unique inherent advantages, such as high-energy efficiency, high-speed modulation capability, low bias currents, easy packaging, and highly integrable structures [22,23]. In particular, VCSELs have demonstrated the ability to generate fast spiking dynamics analogous to those of biological neurons known for their robustness to input noise [24–29]. The controlled activation, inhibition, and communication of these neuronal dynamics has been demonstrated, and recently a single VCSEL device was used to perform spiking pattern recognition and rate coding [24–31]. Thus, photonic spiking VCSELs make suitable candidates for a new future photonic platform for ultrafast energy efficient spiking CNNs.

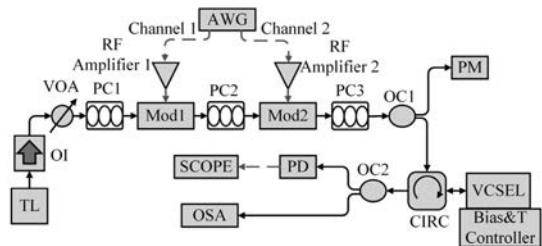
In this work, we use a VCSEL-based photonic approach for binary convolution to demonstrate image gradient magnitude calculation. This delivers an essential portion of the image edge detection functionality used by computer vision and image recognition systems. Here, a single VCSEL system is developed to solely perform a convolution operation; hence, no VCSEL-based CNN architecture, capable of providing learning and classification capabilities, is discussed in this work. The rest of the paper is organized as follows. Section 2 is devoted to the experimental setup of this work for the demonstration of all-optical binary convolution with a spiking VCSEL neuron and the theoretical model used to predict the response of the system. In Section 3, convolutional results are analyzed before the full calculation of image gradient magnitudes is performed both experimentally and theoretically. Finally, Section 4 summarizes the conclusions of this work.

## 2. EXPERIMENTAL SETUP AND THEORETICAL MODEL

We present here the experimental arrangement and theoretical model of the all-optical binary convolution system based on a photonic spiking VCSEL neuron. In this work, we set a source digital image and a kernel as the two inputs of the convolution system. The value of any one pixel in the source image or kernel is limited to 0 or 1.

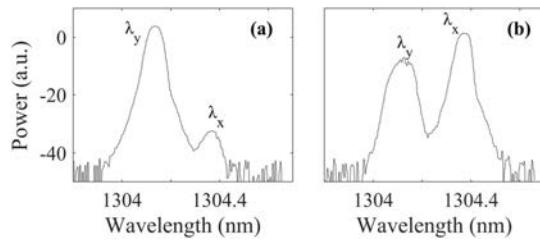
### A. Experimental Setup

Figure 1 shows the schematic diagram of the fiber-optic experimental setup. Two separate electrical signals are generated with a high-bandwidth arbitrary waveform generator (AWG, Keysight M8190a) representing the source image and the kernel used for the convolution process, respectively. These electrical signals (from Channels 1 and 2 of the AWG) are individually amplified by RF Amplifiers 1 and 2 before they are fed into two 10 GHz Mach-Zehnder (MZ) intensity modulators (Mod1 and Mod2) to encode the source image and



**Fig. 1.** Experimental setup of the binary convolution system based on a single VCSEL. TL, tunable laser; OI, optical isolator; VOA, variable optical attenuator; PC1, PC2, and PC3, polarization controllers; AWG, arbitrary waveform generator; Mod1, Mod2, Mach-Zehnder modulators; OC1, OC2, optical couplers; CIRC, circulator; Bias & T Controller, bias and temperature controller; PD, photodetector; PM, power meter; SCOPE, oscilloscope; OSA, optical spectrum analyzer.

kernel into an external optical signal. Compared with two paths of external optical injection, one external optical path, connecting two modulators, makes the injection locking of the VCSEL device easier. It also allows an easy approach to generate (in a single optical path) the required multi-level optical input signal. Additionally, using less optical devices reduces the energy consumption of the external optical path and reduces the cost of the photonic spiking neural network. The external optical signal is generated by a 1300 nm tunable laser (TL, Santec TLS-210 V). An optical isolator (OI) is included after the TL to avoid unwanted light reflections that might lead to spurious results. A variable optical attenuator (VOA) is used after the OI to adjust the strength of the light signal from the TL. The polarization of the optical signal from the TL is adjusted using three polarization controllers (PC1, PC2, and PC3), where PC1 and PC2 are specifically used to match the polarization of the optical signal to that which maximizes the performance of the two modulators, encoding, respectively, the image (Mod1) and the kernel (Mod2) information into the optical path. PC3 is used to adjust the final polarization of the encoded optical signal such that it matches the polarization of the targeted VCSEL mode. A 50:50 optical coupler (OC1) is used to split the light signal into two paths. The first one is connected to a power meter (PM) to monitor the input strength, whilst the second one is directly injected into a commercially available 1300 nm VCSEL through an optical circulator (CIRC). The output of the VCSEL, acting as a spiking optical neuron, is sent to an 8 GHz real-time oscilloscope (SCOPE, Rohde & Schwarz RTP) and an optical spectrum analyzer (OSA, Anritsu MS9710C) for analysis. The VCSEL was kept at a constant temperature of 293 K with an applied bias current of 6.5 mA (the lasing threshold current of the VCSEL was  $I_{th} = 2.96$  mA at 293 K). The optical spectrum of the free-running VCSEL is shown in Fig. 2(a), where the two lasing peaks correspond to the two orthogonal polarizations of the fundamental transverse mode of the device. We refer to the main lasing mode as the parallel polarized mode (or Y-polarized mode, YP mode,  $\lambda_y$ ) and to the subsidiary mode as the orthogonally polarized mode (or X-polarized mode, XP mode,  $\lambda_x$ ). Figure 2(b) shows, in turn, the optical spectrum of the



**Fig. 2.** (a) Optical spectrum of free-running VCSEL used in the experiment. (b) Optical spectrum of the VCSEL subject to constant optical injection. Two polarization modes of VCSELs are referred to as  $\lambda_y$  (parallel) and  $\lambda_x$  (orthogonal).

1300 nm VCSEL device in the spiking regime, as it is subject to optical injection into the orthogonally polarized mode of the device. Upon injection of the external optical signal into the XP mode of the device, the XP mode becomes the dominant mode, whilst the YP mode becomes attenuated. The frequency detuning between the external optically injected signal and the XP mode of the VCSEL was equal to -5.64 GHz. The power of the optically injected signal was 127  $\mu$ W.

### B. Theoretical Model

We use an extension of the well-known spin-flip model (SFM) to model the operation of the VCSEL acting as a spiking optical neuron. In our formulation, we add extra terms to the model's equations to account for the source image and kernel inputs. The rate equations can be described as follows [26,27]:

$$\begin{aligned} \frac{dE_{x,y}}{dt} = & -(k \pm \gamma_a)E_{x,y} - i(k\alpha \pm \gamma_p)E_{x,y} \\ & + k(1 + i\alpha)(NE_{x,y} \pm inE_{y,x}) \\ & + k_{\text{inj}}[E_{\text{inj}x1}(t) + E_{\text{inj}x2}(t)]e^{i\Delta\omega_x t} + F_{x,y}, \end{aligned} \quad (1)$$

$$\frac{dN}{dt} = -\gamma_N[N(1 + |E_x|^2 + |E_y|^2) - \mu + in(E_y E_x^* - E_x E_y^*)], \quad (2)$$

$$\frac{dn}{dt} = -\gamma_s n - \gamma_N[n(|E_x|^2 + |E_y|^2) + iN(E_y E_x^* - E_x E_y^*)], \quad (3)$$

where the subscripts  $x, y$  represent the XP and YP modes of the VCSEL, respectively.  $E_{x,y}$  is the slowly varying complex amplitude of the field in the XP and YP modes.  $N$  is the total carrier inversion between conduction and valence bands.  $n$  is the difference between carrier inversions with opposite spins.  $k$  denotes the field decay rate.  $\gamma_a$  and  $\gamma_p$  are the linear dichroism and the birefringence rate, respectively.  $\alpha$  is the linewidth enhancement factor.  $\gamma_N$  is the decay rate of  $N$ .  $\gamma_s$  is the spin-flip rate.  $\mu$  represents the normalized pump current.  $k_{\text{inj}}$  is the injected strength, and  $E_{\text{inj}x1}$  and  $E_{\text{inj}x2}$  indicate, respectively, the source image and kernel inputs.  $\Delta\omega_x$  is defined as  $\Delta\omega_x = \omega_{\text{inj}x} - \omega_0$ , where  $\omega_{\text{inj}x}$  is the angular frequency of the externally injected light in the XP mode, and  $\omega_0 = (\omega_x + \omega_y)/2$  is the center frequency between the XP and YP modes with

$\omega_x = \omega_0 + \alpha\gamma_a - \gamma_p$  and  $\omega_y = \omega_0 - \alpha\gamma_a + \gamma_p$ . The frequency detuning between the externally injected signal and the XP mode is set as  $\Delta f_x = f_{\text{inj}x} - f_x$ . Hence, in Eq. (1),  $\Delta\omega_x = 2\pi\Delta f_x + \alpha\gamma_a - \gamma_p$ .  $F_{x,y}$  are the spontaneous emission noise terms, which can be written as

$$F_x = \sqrt{\frac{\beta_{\text{sp}}\gamma_N}{2}} \left( \sqrt{N+n}\xi_1 + \sqrt{N-n}\xi_2 \right), \quad (4)$$

$$F_y = -i\sqrt{\frac{\beta_{\text{sp}}\gamma_N}{2}} \left( \sqrt{N+n}\xi_1 - \sqrt{N-n}\xi_2 \right), \quad (5)$$

where  $\beta_{\text{sp}}$  is the strength of the spontaneous emission, and  $\xi_1$  and  $\xi_2$  are independent complex Gaussian white noise terms of zero mean and a unit variance. We numerically solve Eqs. (1)–(4) using the fourth-order Runge–Kutta method. The parameter values configured for the 1300 nm VCSEL are as follows [27]:  $k = 185 \text{ ns}^{-1}$ ,  $\gamma_a = 2 \text{ ns}^{-1}$ ,  $\gamma_p = 128 \text{ ns}^{-1}$ ,  $\alpha = 2$ ,  $\gamma_N = 0.5 \text{ ns}^{-1}$ ,  $\gamma_s = 110 \text{ ns}^{-1}$ ,  $\beta_{\text{sp}} = 10^{-6}$ , and  $k_{\text{inj}} = 125 \text{ ns}^{-1}$ . With these parameters, the YP mode is the main lasing mode, and the XP mode is the subsidiary mode, as in Fig. 2(a).

## 3. EXPERIMENTAL AND NUMERICAL RESULTS

In this section, we firstly provide an experimental proof-of-concept demonstration of all-optical binary convolution with a spiking VCSEL neuron. We then calculate the image gradient magnitudes from a basic “Square” source image and a complex “Horse head” source image by means of all-optical binary convolution. Simulation results on the binary convolution and the calculation of image gradient magnitudes are also presented using a “Horse” source image from the latest version of the Berkeley Segmentation Data Set [32]. Finally, the robustness of our binary convolution system is also tested numerically by adding noise to the source image and kernel inputs.

### A. Experimental Results

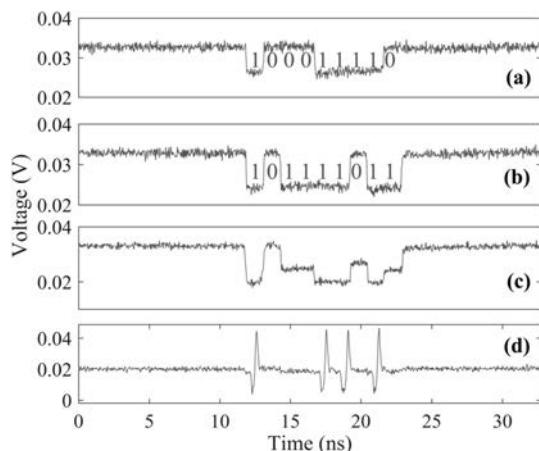
Figure 3 shows an example of a binary two-dimensional (2D) convolution calculation, where a  $3 \times 3$  submatrix (9 pixels) from a source image and a kernel are element-wise multiplied, and the subsequent values of the multiplication are summed. In our experiment, we temporally encoded each pixel of the source image and the kernel inputs using rectangular pulses. Pixels of value “1” were optically encoded using intensity modulated power drops in the TL’s light (via MZ modulators, Mod 1 and Mod 2), whereas pixels of value “0” produced no intensity modulation in the TL’s light. The duration of each rectangular pulse encoding a pixel was set to 1.5 ns to match the refractory period of the experimentally measured spiking dynamics from the VCSEL neuron [17]. The experimental optical realization of the binary convolution example provided in Fig. 3 is depicted graphically in Fig. 4. Figures 4(a) and 4(b) plot, respectively, the temporally encoded 9 pixel ( $3 \times 3$ ) image submatrix and kernel inputs generated for the example given in Fig. 3. Given that the optically encoded source image and kernel inputs were injected into the VCSEL synchronously, we delayed the kernel input such that its modulation (in Mod 2) occurred on top of the corresponding modulated image

$$\begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 1 \\ \hline 1 & 1 & 0 \\ \hline \end{array} \otimes \begin{array}{|c|c|c|} \hline 1 & 0 & 1 \\ \hline 1 & 1 & 1 \\ \hline 0 & 1 & 1 \\ \hline \end{array} = 4$$

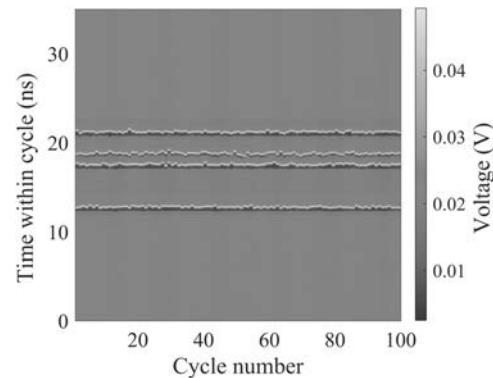
Image                    Kernel

**Fig. 3.** Example of a single step during a 2D binary convolution operation. During this step, a Hadamard (element-wise) product is calculated for a submatrix of the image and the kernel, and all of the values in the multiplication result are summed up to obtain a single value.

input (from Mod 1). We introduced a delay time in the kernel input (directly using the AWG) equal to the time required for a light pulse to travel from Mod 1 to Mod 2. Figure 4(c) shows the optical signal measured after Mod 2 in the setup, combining in a single input line the temporal image and kernel information given in Figs. 4(a) and 4(b). This signal, which was injected into the VCSEL neuron to perform the binary convolution had three different levels (low, medium, and high) depending on the specific pixel values in the image and kernel at a given instance. We control the conditions of the injected signal [in Fig. 4(c)] in such a way that the medium and high input levels injection lock the VCSEL to the external signal, delivering a constant stable temporal output. The lowest input level brings the VCSEL out of the injection locking and into a dynamical region, where the device produces fast spiking dynamical responses [24]. Figure 4(d) shows the experimentally measured time series at the VCSEL neuron's output, yielding stable or spiking outputs depending on the input intensity levels [from Fig. 4(c)]. Importantly, Fig. 4(d) shows that the number of spikes fired by the VCSEL neuron directly provides the result of the binary convolution. It can be seen in Fig. 4(d) that four fast (<100 ps long) spikes are fired by the VCSEL neuron, the same result as that of the binary convolution example in Fig. 3.



**Fig. 4.** Experimental convolution operation. (a) Inputs of Channel 1 (image in Fig. 3). (b) Inputs of Channel 2 (kernel in Fig. 3). (c) Inputs of VCSEL. (d) Outputs of VCSEL (the results of convolution).



**Fig. 5.** Temporal map of 100 superimposed consecutive convolutional results measured experimentally at the output of spiking VCSEL neuron.

Figure 5 shows a temporal map [17] merging in a single plot 100 superimposed consecutive convolutional outputs from the photonic spiking VCSEL neuron. The image and kernel inputs and the experimental conditions are the same as those shown in Fig. 4. Spike events are depicted in yellow in the color map of Fig. 5, and steady state responses appear in light blue. Figure 5 clearly shows that binary convolutional results to 100 consecutive inputs remain the same, producing, in all 100 cases, four separate spiking responses at the VCSEL's output. The optical binary convolutional results obtained with the spiking VCSEL neuron are, therefore, consistent and reproducible. This proof-of-concept result obtained with a spiking VCSEL highlights a new, controllable way to perform convolution operations for information and image processing tasks.

## B. Calculation of Image Gradient Magnitudes

In this section, the image gradient magnitude, critical to image edge detection, is calculated using our approach based on a single spiking VCSEL neuron and optical binary convolution. The image gradient magnitude  $G(x)$  of a given pixel  $x$  is calculated using the following equations [33]:

$$G(x) = \sqrt{G_X(x)^2 + G_Y(x)^2}, \quad (6)$$

$$G_X(x) = [B(x) \otimes B_X^+] - [B(x) \otimes B_X^-], \quad (7)$$

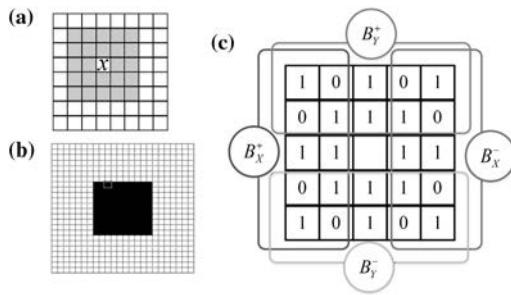
$$G_Y(x) = [B(x) \otimes B_Y^+] - [B(x) \otimes B_Y^-]. \quad (8)$$

Four binary convolutions, i.e.,  $B(x) \otimes B_{X,Y}^\pm$ , are used in  $G_X(x)$  and  $G_Y(x)$ .  $B(x) = \sum_{p=0}^{N-1} s(i_p, i_x) \cdot 2^p$  is the  $N$  bit local binary pattern descriptor of pixel  $x$ .  $i_x$  is the central pixel intensity, and  $i_p$  is the intensity of the  $p$ th neighbor of  $x$  in the source pattern. The comparison operator is defined as

$$s(i_p, i_x) = \begin{cases} 1 & \text{if } |i_p - i_x| > T_x, \\ 0 & \text{otherwise} \end{cases}, \quad (9)$$

where  $T_x = \frac{1}{4}i_x + 20$  and  $N = 5 \times 5 - 1$ .

The range of the local binary pattern descriptor of a pixel is presented in gray color in Fig. 6(a). In Fig. 6(b), a "Square" source image is made up of a solid black  $10 \times 10$  pixels square



**Fig. 6.** (a) Gray color: range of the local binary pattern descriptor of pixels. (b) A  $24 \times 24$  pixels “Square” source image. The red highlight indicates a given pixel in the image. (c) The four convolutions ( $B_X^+$ ,  $B_X^-$ ,  $B_Y^+$ , and  $B_Y^-$ ) of the  $5 \times 5$  binary pattern. Bits that fall outside the highlighted areas for a given string are set to zero.

on a  $24 \times 24$  pixels white background. In the grayscale image, the intensities of white and black pixels are 255 and 0, respectively. For example, the intensity of the red-highlighted pixel  $x$  in Fig. 6(b) is  $i_x = 255$ . We arrange and serialize the pixels in the range of local binary pattern descriptors by columns. The first neighbor pixel intensity is  $i_1 = 0$ ; hence, according to Eq. (9),  $s(i_1, i_x) = 1$ . The third neighbor is  $i_3 = 255$ ; hence,  $s(i_3, i_x) = 0$ .  $B(x)$  can be calculated for the red-highlighted pixel in Fig. 6(b) as follows:

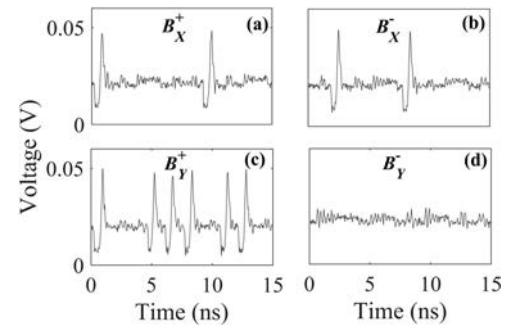
$$B(x) = \begin{bmatrix} s(i_1, i_x) & - & - & - & s(i_{20}, i_x) \\ s(i_2, i_x) & - & - & - & s(i_{21}, i_x) \\ s(i_3, i_x) & - & x & - & s(i_{22}, i_x) \\ s(i_4, i_x) & - & - & - & s(i_{23}, i_x) \\ s(i_5, i_x) & - & - & - & s(i_{24}, i_x) \end{bmatrix} \\ = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & x & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (10)$$

For the red-highlighted pixel  $x$  in Fig. 6(b), “1” in  $B(x)$  corresponds to a white pixel, and “0” corresponds to a black pixel in the source image.

In Eqs. (7) and (8),  $B_X^+$ ,  $B_X^-$ ,  $B_Y^+$ , and  $B_Y^-$  are the four kernels that are adopted as in Ref. [33]. Figure 6(c) shows the areas of the four different kernels. Pixels that fall outside of the highlighted areas in Fig. 6(c) for a given string are set to zero. For example,

$$B_X^+ = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (11)$$

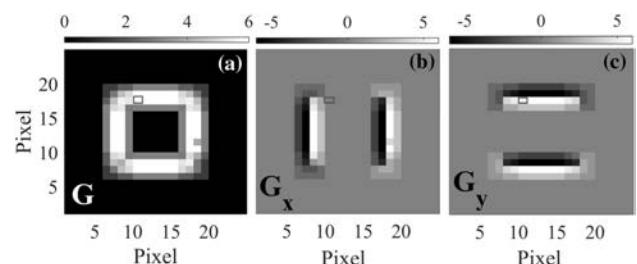
We arrange and serialize the pixels of  $B(x)$  and the four kernels by columns. For example, the string of  $B(x)$  is  $[1, 1, 0, 0, 1, 1, 0, 0, 0, \dots]$ , and the string of  $B_X^+$  is  $[1, 0, 1, 0, 1, 0, 1, 1, 1, 0, \dots]$ . We studied experimentally the response of the VCSEL



**Fig. 7.** Four convolutional results with four highlighted area kernels for one pixel, which has red box in Fig. 6.

neuron under the injection of the “Square” source image and kernel operators included in Figs. 6(b) and 6(c). Specifically, Fig. 7 showcases the experimentally recorded results at the VCSEL output for each kernel when operating on the red-highlighted pixel in Fig. 6(b). It can be seen in Fig. 7(a) that fast [sub-nanosecond (ns)] spikes are only triggered by the 1st and 7th pixels. Therefore, the convolutional result for  $B(x) \otimes B_X^+$  is two, as was expected. Here, the convolutional result is measured offline, where the number of spiking responses is counted using software. This could be achieved in future realizations experimentally using electronic or photonic spike/photon counting hardware. Similarly, from Figs. 7(b)–7(d) we can see that 2, 6, and 0 sub-ns spikes are elicited at the VCSEL’s output for kernels  $B_X^-$ ,  $B_Y^+$ , and  $B_Y^-$ , respectively. Using the experimental results measured from the spiking VCSEL neuron, we calculate off-line  $G_X(x)$ ,  $G_Y(x)$ , and  $G(x)$  to determine the image gradient magnitude. Based on the experimentally measured results in Figs. 7(a)–7(d),  $G_X(x)$ ,  $G_Y(x)$ , and  $G(x)$  are 0, 6, and 6, respectively, using Eqs. (7)–(9).

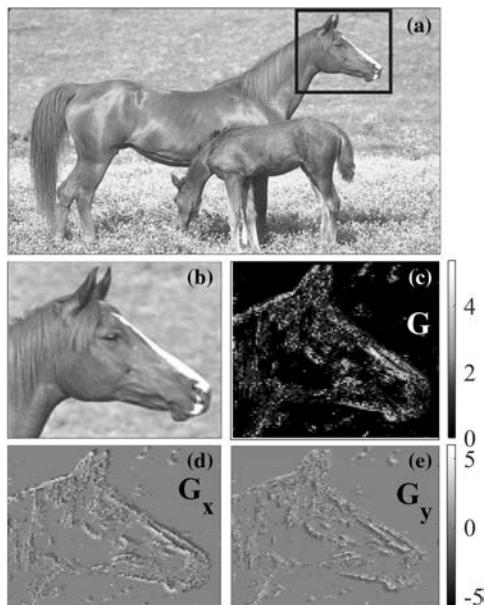
The experimental process in Fig. 7 is repeated consecutively for every single pixel in the “Square” source image [Fig. 6(b)] to calculate their image gradient magnitudes. The latter are used to build the reconstructed image in Fig. 8(a), providing a gradient map for the “Square” source image. Figure 8(a) clearly reveals a “hollow” square shape in the experimentally produced gradient map, hence, detecting all edge features of the source image. Here, the corner and edge pixels are omitted [34]. In Fig. 8(a), the pixels with a gradient magnitude  $G(x) > 3$  can be selected to thin the response and reveal the true edges



**Fig. 8.** Gradient maps of the “Square” source image. Visualizations of (a)  $G$ , (b)  $G_X$ , and (c)  $G_Y$  maps of the “Square” source image based on the optical binary convolution performed by the VCSEL neuron.

of the “Square” [33,35]. Additionally, Figs. 8(b) and 8(c) plot separately the reconstructed images using the obtained values for  $G_X(x)$  and  $G_Y(x)$  from the experimentally measured time series at the VCSEL neuron’s output. Figures 8(b) and 8(c) reveal that both vertical and horizontal lines can be individually detected from the source image in Fig. 6(b) using, respectively, the magnitudes  $G_X(x)$  and  $G_Y(x)$ . For one pixel  $x$ , a total of 15 ns is required to process each of the four binary convolutions, as shown in Fig. 7. Hence, 60 ns (15 ns  $\times$  4) is needed for the binary convolutions of one pixel with our single VCSEL system. The time required for binary convolution is, therefore, dependent on the number of pixels in the image. Considering that the optical power of the VCSEL is on average equal to  $\sim 500 \mu\text{W}$ , for the different operation conditions used in this work, we can estimate the energy consumption of the binary convolution for one pixel as 30 pJ ( $0.5 \text{ mW} \times 60 \text{ ns}$ ) in our system. In the future, the binary convolution operation could be achieved using multiple devices integrated in a VCSEL array simultaneously. Such a new architecture would increase the speed of the convolution operation, obviously at the expense of increasing the system’s complexity. This work therefore provides a low complexity, reduced energy consumption, and fast hardware approach for photonic binary convolution for novel light-enabled image processing functionalities.

To further investigate our experimental system, we focused on demonstrating the achievement of gradient maps from a complex source image using the all-optical binary convolution of this work, as seen in Fig. 9. For this purpose, we selected as a source image for our VCSEL-based binary convolution system a complex “Horse head” image [Fig. 9(b)]. This is a  $100 \times 105$



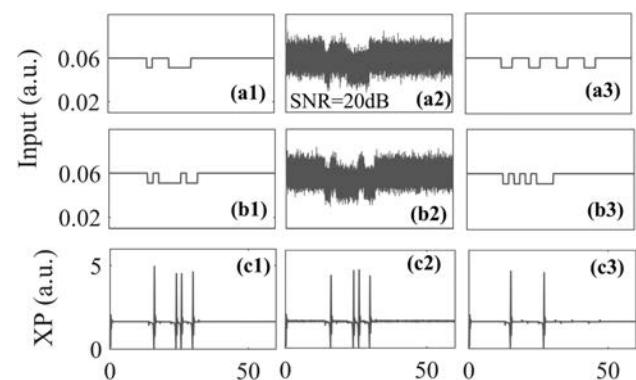
**Fig. 9.** “Horse head” image and the gradient maps of the “Horse head” image. (a) Source “Horse” image. The blue box indicates the “Horse Head” image used for analysis in (b). Visualizations of the (c)  $G$ , (d)  $G_X$ , and (e)  $G_Y$  maps of the “Horse head” image obtained from the optical binary convolution performed with the VCSEL neuron.

pixels portion of the “Horse” image from the Berkeley Segmentation Data Set [32] [also included in Fig. 9(a)]. The color image was converted to grayscale before we applied the same experimental methods used previously to obtain the results included in Fig. 8 above. The values of  $G(x)$ ,  $G_X(x)$ , and  $G_Y(x)$  experimentally achieved for the complex “Horse head” image [Fig. 9(b)] are shown in Figs. 9(c)–9(e), respectively. These gradient maps reveal the successful detection of the edge features in this complex image, hence permitting the successful recreation of the outline and shape of the horse head. This effectively demonstrates that the reported all-optical binary convolution technique with a VCSEL neuron is also suitable for complex high-resolution source images.

### C. Numerical Results

In this section, binary convolution based on a single VCSEL neuron is performed numerically. The robustness of the system to perform all-optical binary convolution under noisy inputs and for larger kernels is investigated. Finally, the calculation of image gradient magnitudes with our photonic approach using a single VCSEL neuron is presented numerically using the “Horse” image from the latest version of the Berkeley Segmentation Data Set [32].

The binary convolution example given in Fig. 3 and experimentally performed with the VCSEL neuron (see Fig. 4) is numerically simulated using the SFM model in Figs. 10(a1)–10(c1). Pixels of value “1” are numerically implemented using power drop pulses with a strength  $K_p = 0.852$  ( $K_p$  = pulse power/constant power) and a duration of 1.5 ns (as in the experimental demonstration). The frequency detuning between the externally injected signal and the XP mode in the VCSEL model is set to  $-3.66 \text{ GHz}$ . Figures 10(a1)–10(c1) plot the numerically obtained results for the all-optical binary convolution with a VCSEL neuron. Specifically, Figs. 10(a1) and 10(b1) plot, respectively, the time series for the temporally encoded image [Fig. 10(a1)] and kernel [Fig. 10(b1)] inputs, whilst Fig. 10(c1) plots the numerically calculated output from the VCSEL neuron. The latter clearly shows that the simulation successfully reproduces the



**Fig. 10.** (a1)–(a3) Inputs of Channel 1 (image in Fig. 3). (b1)–(b3) Inputs of Channel 2 (kernel in Fig. 3). (c1)–(c3) VCSEL neuron’s output. (a1)–(c1) Convolutional operation in the VCSEL neuron without noise. (a2)–(c2) Convolutional operation in the VCSEL neuron with added input noise of  $\text{SNR} = 20 \text{ dB}$ . (a3)–(c3) Convolution operation with a  $5 \times 5$  pixels kernel.

outcome of the experimental all-optical binary convolution [see Fig. 4(d)], where four spikes are elicited by the VCSEL. This excellent agreement between the modeled results and the experimental findings gives us confidence to test the robustness of the photonic binary convolution system under the injection of inputs with added noise. To study this aspect, we model the response of the VCSEL binary convolutional system under the injection of noisy inputs with a configured SNR = 20 dB [see results in Figs. 10(a2) and 10(b2)]. Specifically, Fig. 10(c2) shows that the exact same response is obtained from the VCSEL neuron as compared to the case with no added noise in Fig. 10(c1). This outlines the robustness to noise of the proposed all-optical VCSEL convolutional system. Additionally, the numerical convolution with a larger  $5 \times 5$  pixels kernel is tested numerically in Figs. 10(a3)–10(c3) using Eq. (10) and Eq. (11) as inputs. Figure 10(c3) shows that the modeled convolutional result obtained from the VCSEL neuron also produces two fast spike events, hence yielding the exact same outcome as obtained experimentally in Fig. 7(a). We can therefore deduce that the convolution results that can be obtained with our VCSEL-neuron-based approach are not limited by the dimension of the kernel operators or the resolution of the image.

Figure 11 shows the numerically calculated gradient maps obtained with a spiking VCSEL neuron for the “Horse” source image [32] with a resolution of  $481 \times 321$  pixels [Figs. 11(a) and 9(a)]. Figures 11(b)–11(d) show the calculated gradient maps for  $G(x)$ ,  $G_X(x)$ , and  $G_Y(x)$ , respectively. These were obtained using the  $5 \times 5$  kernel introduced in the experimental study of the “Square” source image (see Figs. 6–8). It can be seen that the numerical simulation successfully reveals the image edge information through the gradient magnitude  $G(x)$ , as seen in Fig. 11(b), as well as the individual horizontal and vertical edge features of the source image through  $G_X(x)$  and  $G_Y(x)$ , as seen in Figs. 11(c) and 11(d), respectively. These results, showing good overall agreement with the experimental findings of Fig. 9, therefore numerically validate that the gradient magnitude can be successfully calculated with a

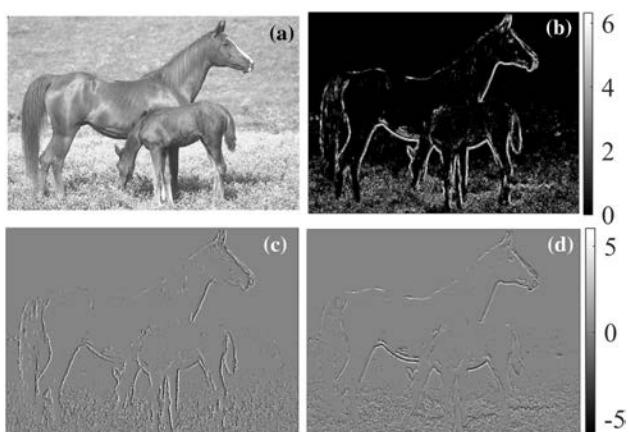
photonic spiking VCSEL neuron, irrespective of the image dimensionality.

Optical binary convolution can be used in systems where simplified convolutional operations, with binary inputs (still able to provide high-performance accuracy), provide other key advantages in terms of increased operation speed, lowered energy consumption, and reduced hardware requirements. This is the case of the system reported in this work, using an extremely hardware friendly implementation of a single VCSEL to perform high-speed and low energy ( $<\mu\text{J}/\text{spike}$ ) image edge-feature detection. Besides, in our approach, the results of the binary convolution are output in an optical spiking representation, providing unique advantages in terms of robustness to noise and high precision of convolutional results. This spiking representation therefore enables our platform to successfully perform with noisy optical and electronic signals. Whilst other recent works have recently reported complex systems for optical convolution operation using temporally modulated inputs and weights for image processing tasks showing excellent performance [36], our technique benefits from an extremely simple architecture using just one off-the-shelf, inexpensive 1300 nm VCSEL to perform the binary convolution operation for image edge-feature detection. Our approach combining a VCSEL-based spiking photonic neuron with time multiplexing is able to deliver the operation of a full neuronal layer, where each 1.5-ns-long time slot operates in fact as a virtual neuron (or node) processing specific image pixel information. This offers great promise for future implementation of interconnected VCSEL-based neuronal network architectures for image processing tasks of increased complexity (e.g., image classification) and using neuron-like spiking signals to operate.

The utilization of binary convolution in the calculation of image gradient maps has been reported to outperform the alternative Canny implementation [35] of image gradient maps convolution, in the Intel i7 mobile processor [33]. In that report, the frequency of the binary gradient-based edge detector was 4.7 Hz, while the Canny convolution approach was found to operate at 0.5 Hz. This indicates that binary convolution can be performed at speeds faster than alternative convolution approaches. Additionally, mobile processors operate with powers of several watts (for example, the Intel i7 has a power of 15 W) [37], whilst VCSELs, such as the one used in this work, provide low power performance typically at milliwatt (mW) and sub-mW power levels. Hence, the energy consumption for the calculation of image gradient maps obtained with our VCSEL-based optical binary convolution system can be significantly more energy efficient, as well as yield faster operation speeds, than the performance achieved with traditional or binary convolution methods in digital processors.

#### 4. CONCLUSION

In this work, we proposed and investigated experimentally and numerically an all-optical binary convolution system using a VCSEL operating as a photonic spiking neuron. The inputs (image and kernel) are encoded temporally using fast rectangular pulses (1.5-ns-long) and optically injected into the VCSEL neuron. The latter's optical output directly provides the results of the convolution in the number of (sub-ns long) spikes fired.



**Fig. 11.** “Horse” image and gradient maps of the “Horse” image. (a) “Horse” image. Visualizations of (b)  $G$ , (c)  $G_X$ , and (d)  $G_Y$  maps of “Horse” image based on the numerical optical binary convolution in VCSEL.

In addition to performing all-optical binary convolution, we demonstrated experimentally and numerically the ability of the proposed system to calculate the image gradient magnitudes from digital source images. This feature was successfully used to identify key edge features from a source image as well as its separate horizontal and vertical components. Furthermore, we investigated numerically the robustness of the proposed VCSEL-based convolutional system to input noise. This simple system, using a single commercially available VCSEL operating at the key telecom wavelength of 1300 nm, offers a novel photonic solution to binary convolution with the advantage of being highly energy efficient and hardware friendly. This opens exciting prospects for a new photonic spiking platform for future optical binary spiking CNNs. Furthermore, the high-speed, low cost, and neuronal functionalities of these photonic spiking systems hold promise for numerous processing tasks expanding into fields such as computer vision and artificial intelligence.

**Funding.** UKRI Turing AI Acceleration Fellowships Programme (EP/V025198/1); Office of Naval Research Global (ONRGNICOP-N62909-18-1-2027); European Commission (828841-ChipAI-H2020-FETOPEN2018-2020); UK's EPSRC Doctoral Training Partnership (EP/N509760); National Natural Science Foundation of China (61674119, 61974177); China Scholarship Council.

**Acknowledgment.** We thank Prof. T. Ackemann and Prof. A. Kemp (University of Strathclyde) for lending some of the equipment used in this work. All data underpinning this publication are openly available from the University of Strathclyde KnowledgeBase at <https://doi.org/10.15129/51af27bc-8cc2-46a7-9088-1aad87e4340c>.

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- O. Abdel-Hamid, A. R. Mohamed, H. Jiang, and G. Penn, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (2012), pp. 4277–4280.
- J. Fu, H. Zheng, and T. Mei, "Look closer to see better: recurrent attention convolutional neural network for fine-grained image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4438–4446.
- K. Gopalakrishnan, S. K. Khaitan, A. Choudhary, and A. Agrawal, "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection," *Constr. Build. Mater.* **157**, 322–330 (2017).
- P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *International Conference on Document Analysis and Recognition (ICDAR)* (2003), pp. 958–963.
- C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 1915–1929 (2013).
- L. Cavigelli, M. Magno, and L. Benini, "Accelerating real-time embedded scene labeling with convolutional networks," in *Proceedings of the 52nd Annual Design Automation Conference* (2015), paper 108.
- M. Rastegari, V. Ordonez, J. Redmon, and A. Farhadi, "Xnor-net: ImageNet classification using binary convolutional neural networks," in *European Conference on Computer Vision (ECCV)* (2016), pp. 525–542.
- F. Juefei-Xu, V. N. Boddeti, and M. Savvides, "Local binary convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 19–28.
- X. Lin, C. Zhao, and W. Pan, "Towards accurate binary convolutional neural network," in *Advances in Neural Information Processing Systems (NIPS)* (2017), pp. 345–353.
- M. Courbariaux, Y. Bengio, and J.-P. David, "BinaryConnect: training deep neural networks with binary weights during propagations," in *Advances in Neural Information Processing Systems (NIPS)* (2015), pp. 3105–3113.
- M. Courbariaux, I. Hubara, D. Soudry, R. El-Yaniv, and Y. Bengio, "Binarized neural networks: training deep neural networks with weights and activations constrained to +1 or -1," *arXiv:1602.02830* (2016).
- M. Turconi, B. Garbin, M. Feyereisen, M. Giudici, and S. Barland, "Control of excitable pulses in an injection-locked semiconductor laser," *Phys. Rev. E* **88**, 022923 (2013).
- P. R. Prucnal, B. J. Shastri, T. F. de Lima, M. A. Nahmias, and A. N. Tait, "Recent progress in semiconductor excitable lasers for photonic spike processing," *Adv. Opt. Photon.* **8**, 228–299 (2016).
- S. Xiang, Y. Zhang, J. Gong, X. Guo, L. Lin, and Y. Hao, "STDP-based unsupervised spike pattern learning in a photonic spiking neural network with VCSELs and VCSOAs," *IEEE J. Sel. Top. Quantum Electron.* **25**, 1700109 (2019).
- Y. Zhang, S. Xiang, X. Guo, A. Wen, and Y. Hao, "All-optical inhibitory dynamics in photonic neuron based on polarization mode competition in a VCSEL with an embedded saturable absorber," *Opt. Lett.* **44**, 1548–1551 (2019).
- J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–214 (2019).
- J. Robertson, E. Wade, Y. Kopp, J. Bueno, and A. Hurtado, "Toward neuromorphic photonic networks of ultrafast spiking laser neurons," *IEEE J. Sel. Top. Quantum Electron.* **26**, 7700715 (2019).
- A. Mehrabian, Y. Al-Kabani, V. J. Sorger, and T. El-Ghazawi, "PCNNA: a photonic convolutional neural network accelerator," in *IEEE International System-on-Chip Conference (SOCC)* (2018), pp. 169–173.
- H. Bagherian, S. Skirlo, Y. Shen, H. Meng, V. Cepric, and M. Soljacic, "On-chip optical convolutional neural networks," *arXiv:1808.03303* (2018).
- S. Xu, J. Wang, R. Wang, J. Chen, and W. Zou, "High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays," *Opt. Express* **27**, 19778–19787 (2019).
- S. Xu, J. Wang, and W. Zou, "High-energy-efficiency integrated photonic convolutional neural networks," *arXiv:1910.12635* (2019).
- K. Iga and H. E. Li, *Vertical-Cavity Surface-Emitting Laser Devices* (Springer, 2003).
- R. Michalzik, *VCSELs: Fundamentals, Technology and Applications of Vertical-Cavity Surface-Emitting Lasers*, Springer Series in Optical Sciences (Springer-Verlag, 2013), Vol. **166**.
- A. Hurtado and J. Javaloyes, "Controllable spiking patterns in long-wavelength vertical cavity surface emitting lasers for neuromorphic photonics systems," *Appl. Phys. Lett.* **107**, 241103 (2015).
- B. Garbin, J. Javaloyes, G. Tissoni, and S. Barland, "Topological solitons as addressable phase bits in a driven laser," *Nat. Commun.* **6**, 5915 (2015).
- S. Y. Xiang, A. J. Wen, and W. Pan, "Emulation of spiking response and spiking frequency property in VCSEL-based photonic neuron," *IEEE Photon. J.* **8**, 1504109 (2016).
- T. Deng, J. Robertson, and A. Hurtado, "Controlled propagation of spiking dynamics in vertical-cavity surface-emitting lasers: towards neuromorphic photonic networks," *IEEE J. Sel. Top. Quantum Electron.* **23**, 1800408 (2017).
- J. Robertson, T. Deng, J. Javaloyes, and A. Hurtado, "Controlled inhibition of spiking dynamics in VCSELs for neuromorphic photonics: theory and experiments," *Opt. Lett.* **42**, 1560–1563 (2017).

29. A. Dolcemascolo, B. Garbin, B. Peyce, R. Veltz, and S. Barland, "Resonator neuron and triggering multipulse excitability in laser with injected signal," *Phys. Rev. E* **98**, 062211 (2018).
30. J. Robertson, M. Hejda, J. Bueno, and A. Hurtado, "Ultrafast optical integration and pattern classification for neuromorphic photonics based on spiking VCSEL neurons," *Sci. Rep.* **10**, 1 (2020).
31. M. Hejda, J. Robertson, J. Bueno, and A. Hurtado, "Spike-based information encoding in VCSELs for neuromorphic photonic systems," *J. Phys.* **2**, 044001 (2020).
32. P. Arbelaez, M. Maire, C. Fowlkes, and J. Malik, "Contour detection and hierarchical image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.* **33**, 898–916 (2011).
33. P. L. St-Charles, G. A. Bilodeau, and R. Bergevin, "Fast image gradients using binary feature convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp.1–9.
34. E. Nadernejad, "Edge detection techniques: evaluations and comparisons," *Appl. Math. Sci.* **2**, 1507–1520 (2008).
35. J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 679–698 (1986).
36. R. Hamerly, L. Bernstein, A. Sludds, M. Soljačić, and D. Englund, "Large-scale optical neural networks based on photoelectric multiplication," *Phys. Rev. X* **9**, 021032 (2019).
37. V. Spiliopoulos, G. Keramidas, S. Kaxiras, and K. Efstathiou, "Power-performance adaptation in Inter core i7," in *Proceedings of 2nd Workshop Computer Architecture and Operating System Co-design* (2011), pp. 1–10.

# PHOTONICS Research

## Imaging through unknown scattering media based on physics-informed learning

SHUO ZHU,<sup>1,†</sup> ENLAI GUO,<sup>1,2,†</sup> JIE GU,<sup>1</sup> LIANFA BAI,<sup>1</sup> AND JING HAN<sup>1,3</sup>

<sup>1</sup>Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense, Nanjing University of Science and Technology, Nanjing 210094, China

<sup>2</sup>e-mail: njjustgel@njjust.edu.cn

<sup>3</sup>e-mail: eohj@njjust.edu.cn

Received 2 December 2020; revised 26 January 2021; accepted 21 February 2021; posted 22 February 2021 (Doc. ID 416551); published 15 April 2021

Imaging through scattering media is one of the hotspots in the optical field, and impressive results have been demonstrated via deep learning (DL). However, most of the DL approaches are solely data-driven methods and lack the related physics prior, which results in a limited generalization capability. In this paper, through the effective combination of the speckle-correlation theory and the DL method, we demonstrate a physics-informed learning method in scalable imaging through an unknown thin scattering media, which can achieve high reconstruction fidelity for the sparse objects by training with only one diffuser. The method can solve the inverse problem with more general applicability, which promotes that the objects with different complexity and sparsity can be reconstructed accurately through unknown scattering media, even if the diffusers have different statistical properties. This approach can also extend the field of view (FOV) of traditional speckle-correlation methods. This method gives impetus to the development of scattering imaging in practical scenes and provides an enlightening reference for using DL methods to solve optical problems. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416551>

### 1. INTRODUCTION

Object information is seriously degraded after being modulated by complex media [1,2]. The scattering of light in the diffusion is an established problem considered to be a common phenomenon in our daily life (i.e., seeing through dense fog to obtain the license plate and driver's facial information is crucial for a traffic monitor). Imaging with randomly scattered light is a challenging problem with an urgent requirement in different fields (e.g., astronomical observations through the turbulent atmosphere and biological analysis through active tissue) [3–7]. Conventional imaging methods based on geometric optics cannot work with the disordered light beam under scattering. Benefiting from the great progress of optoelectronic devices and computational techniques, many new imaging methods have been proposed for imaging through scattering media. The typical imaging techniques include the wave-front-shaping methods [8–11], reconstruction using the transmission matrix [12,13], single-pixel imaging methods [14–16], or techniques based on the point spread function (PSF) [17–19]. The methods listed here have made great progress in object reconstruction with invasive prior, and the scattering scenes are relatively stable. Speckle correlation based on the optical memory effect (OME) is an extraordinary method for noninvasive imaging through opaque layers [20] with only one frame of speckle pattern [21,22]. Object recovery based

on speckle correlation methods uses phase retrieval algorithms such as hybrid input-output (HIO) [23], those based on the alternating direction method of multipliers (ADMM) [24], and phase retrieval based on generalized approximate message passing (prGAMP) [25]. The field of view (FOV) of speckle-correlation methods is limited by OME, and the recovery performance is also influenced by the recovery capability of phase retrieval algorithms.

Recently, with the advent of digital technology, big data, and advanced optoelectronic technology, deep learning (DL) has shown great potential in optics and photonics [26,27]. With powerful data mining and mapping capabilities, the data-informed DL methods can extract the key features and build a reliable model in many fields [28]. To date, the DL approach has been successfully applied in digital holography imaging [29–32], Fourier ptychographic imaging [33–36], computational ghost imaging [37,38], superresolution microscopic imaging [39–42], optical tomography imaging [43–45], photon-limited imaging [46,47], three-dimensional (3D) measurements with fringe pattern analysis [48–51], and imaging through scattering media [52–60]. Compared to traditional computational imaging (CI) technology, the learning-heuristic methods can not only solve complex imaging problems, but also can significantly improve the core performance indicators (i.e., spatial resolution, temporal resolution, and sensitivity).

The great progress by DL is indicated by the rapidly increasing number of DL-related publications in photonics journal in the last several years [61]. However, the methods using DL are meeting several challenging problems, such as the choice of the DL framework tending to be empirical and the limited generalization capability.

Based on the nonlinear characteristics of deep neural networks (DNNs), DL methods have good performance in highly ill-posed problems, especially in imaging through random media [52–54,57]. IDiffNet is the first proposed method to reconstruct an object through scattering media via a densely connected DNN. The performance with a different type of training dataset and the loss function are systematically discussed [52]. A hybrid neural network is constructed to see through a thick scattering medium and achieves object restoration exceeding the FOV of OME [53]. The speckle patterns of single-mode fiber and multi-mode fibers are reconstructed and recognized successfully [54]. PDSNet is built to reconstruct complex objects through scattering medium and expands the FOV up to 40 times of the memory effect (ME) [55]. The methods above are mainly focused on a specified diffuser that has a limitation in complex and variable scattering conditions. Therefore, some DL methods have the potential to reconstruct objects through unstable media that mainly use different DNN structures, such as one-to-all with dense blocks, interpretable DL method, generative adversarial network (GAN), or a two-stage framework [57–60]. Li *et al.* [57] first proposed a DL technique to generalize from four different diffusers to more diffusers with raw speckles, which requires the unknown diffusers to have similar statistical properties and the structure of objects to be simple [57]. Almost all the DL methods for imaging through scattering media use speckle patterns directly, and more information might be further excavated with a traditional physical theory. The efficient physics prior can provide an optimized direction for DNN to find the optimal reconstruction solution in different scattering scenes. After being modulated by different diffusers, the scattering light with photons walk randomly brings about the great statistical difference of speckle patterns even with the same one object. Although it has been proven that the DL method, which focuses on DNN structure design, has the generalization capability to reconstruct hidden objects through unknown diffusers, it is still difficult to obtain an accurate object structure under the condition of fewer training diffusers and has a limitation in reconstructing the complex object [57]. At the same time, the generalized diffusers should have similar statistical characteristics. Therefore, in the absence of effective physical constraints and guidance, DL methods can hardly extract universal information from speckle patterns under highly degenerating conditions. Solely data-driven DL methods will lead to the limited generalization capability that the model is over-relying on training data. Thus, to solve the problems of imaging through multicomplex media, combining the scattering theory with DNN is a more efficient method than designing specific DNN structures.

In this paper, with the physics prior of scattering and the support of DL, a physics-informed learning method is proposed for imaging through unknown diffusers. By pre-processing, the data model based on the physics prior can solve the generalization problems in different scattering scenes, which can reduce

the data dependence of the DL model and robustly improve the feature extraction efficiency. The efficient physics prior can provide an optimized direction for DNN to find the optimal reconstruction solution in different scattering scenes. The DL method based on physics prior can help to learn and extract the statistical invariants from different scattering scenes. Instead of training with captured patterns directly, using the DL framework with speckle correlation prior to imaging through different diffusers is technologically reasonable. Employing the physics-informed learning method, scalable imaging through unknown diffusers can be achieved with high reconstruction quality. The scattering degradation of the sparse objects can even be modeled with one ground glass, and imaging through unknown ground glasses even with different statistical characteristics can be achieved. More complex objects (e.g., human faces) can be reconstructed accurately by slightly increasing the number of training diffusers. Meanwhile, it is hard to restore the objects efficiently exceeding the FOV of OME by the traditional speckle-correlation method. Based on the powerful capability in data mining and processing of DNN, the physics-informed learning method can also break through the FOV limitation for scalable imaging. Finally, we demonstrate the physics-informed learning scheme with an experimental dataset and present the quantitative evaluation results with multiple indicators. The results with the statistical average indicator show the accuracy and robustness of our scheme, and reflect the great potential of combining physical knowledge and DL.

## 2. METHODS

### A. Physical Basic

The proposed model must be established to have general applicability for scalable imaging through unknown diffusers, and it is also one of the indispensable conditions to apply this method to practical complex scenes. The wave propagating through an inhomogeneous medium with multiple scattering will generate a fluctuating intensity pattern, and the universal physical law exists in different transmitted modes. The speckle correlation and memory effect in optical wave transmission through disordered media are proposed to observe and analyze the shift-invariant characteristic of speckle patterns [62,63]. The speckle patterns of scattered light through diffusive media are invariant to small tilts or shifts in the incident wavefront of light, and the outgoing light field still retains the information carried by the incoming beam within the range of ME [64]. Therefore, within the scope of ME, the scattering system can be considered as an imaging system with shift-invariant point spread function. The speckle pattern captured by the camera is given by the convolution of the object intensity pattern  $O(x)$  with the PSF ( $S$ ), which can be calculated by

$$I = O * S, \quad (1)$$

where the symbol  $*$  denotes the convolution operator. Using the convolution theorem, the autocorrelation of camera pattern intensity can be defined as

$$I \star I = (O * S) \star (O * S) = (O \star O) * (S \star S), \quad (2)$$

where the  $\star$  is the correlation operator and  $S \star S$  is a sharply peaked function representing the autocorrelation of broadband

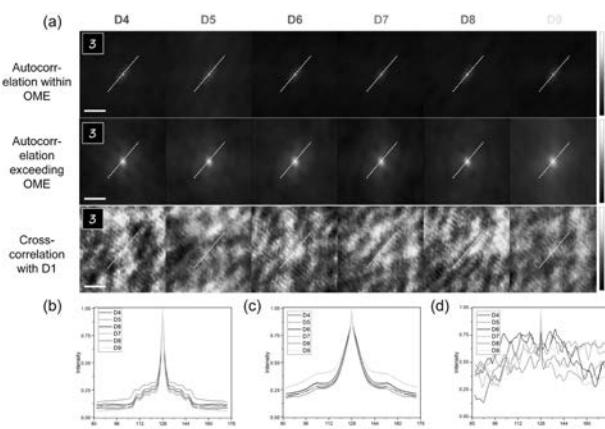
noise. The autocorrelation of the speckle pattern is approximately equal to the autocorrelation of the object hidden behind the scattering media, and the speckle autocorrelation has an additional constant background term  $C$  [21]. Thus, Eq. (2) can be further simplified as

$$I \star I = (O \star O) + C. \quad (3)$$

When the object size exceeds the range of OME, the object can be divided into multiple objects  $O_i$  within the OME scope and  $n$  represents the object distributed in  $n$  different OME ranges (see Appendix A for details). Thus, the autocorrelation distribution of the speckle pattern exceeding OME can be defined as

$$I \star I = \sum_{i=1}^n (O_i \star O_i) + C'. \quad (4)$$

To clarify the universal connection among different scattering scenes, speckle patterns are captured through different diffusers with the same object. In total, we take speckle patterns using nine different ground glasses in the experiment, including six 220 grit diffusers (D1, D2, D3, D4, D5, and D6), one 120-grit diffuser (D7), one 600 grit diffuser (D8) produced by Thorlabs, and one 220 grit diffuser (D9) produced by Edmund. Among these diffusers, D4 to D9 are selected as testing diffusers. As shown in the first row in Fig. 1(a), even when the diffusers have different statistical characteristics and are made by different manufacturers, the autocorrelation of the object has a high degree of similarity with the autocorrelation of the speckle pattern within or exceeding the range of OME, and the difference is reflected in the different background terms. On the other hand, the correlation of different speckles is irregular by calculating the cross-correlation with D1. The autocorrelation of the speckle pattern exceeding OME is also similar among different diffusers to some extent. However, the cross-correlation with different diffusers almost has no



**Fig. 1.** Speckle statistical characteristics analysis of the same object corresponding to different testing diffusers. (a) First row and second row are the speckle autocorrelation of the object within or exceeding the OME range, the third row is the cross-correlation with D1, respectively. (b)–(d) Intensity values of the white dash lines in the first, second, and third rows of (a), respectively. The color bar represents the normalized intensity. Scale bars: 875.52 μm.

similarities, even with similar statistical properties (e.g., D4, D5, and D6).

With the speckle-correlation prior, the statistical invariants of the object through different scattering media can be effectively extracted, which informs the DNN to obtain useful information and reconstruct the object in different scattering scenes. Imaging through different scattering media with speckle-correlation prior can be used as a reference and a heuristic approach to design the DL methods in different optical problems.

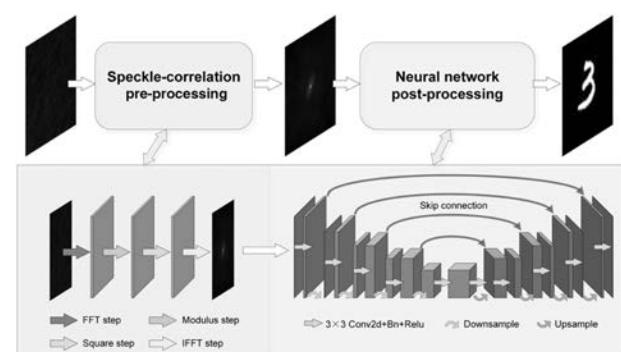
## B. Framework of Physics-Informed Learning

To solve the optical problems by DL methods, it is essential to make full use of optical physics prior. As shown in Fig. 2, the physics-informed learning framework consists of a speckle-correlation pre-processing step and a neural network post-processing step. The pre-processing step is mainly to obtain speckle autocorrelation  $R(x, y)$ , is calculated by an inverse 2D Fourier transform of its energy spectrum, and the mathematical operations can be expressed as

$$R(x, y) = I(x, y) \star I(x, y) = \text{FFT}^{-1}\{\|\text{FFT}\{I(x, y)\}\|^2\}. \quad (5)$$

After the speckle-correlation pre-processing step, the captured speckle pattern is adjusted and refactored, and the next step is post-processing by DNN to reconstruct the hidden object. By adding the speckle-correlation theory, the imaging model can make full use of the advantages of the neural network. The DL model is a simple convolutional neural network (CNN) of the U-Net type [65]. Comparing a specially designed DNN structure, the physics-informed learning method can achieve better imaging results with a simple U-Net without any other tricks.

In our experiments, multiple objects datasets with different levels of complexity are used to reconstruct through different diffusers, such as the modified National Institute of Standards and Technology (MNIST) dataset [66] and FEI face dataset [67]. An equilibrium constraint loss function is important for the training process, and we design a combination loss that includes negative Pearson correlation coefficient (NPCC) loss and mean square error (MSE). The Pearson correlation coefficient is an index used to evaluate the similarity between two variables, and the calculated value is distributed from -1 to 1. A negative value represents a negative correlation,



**Fig. 2.** Schematic of the physics-informed learning method for scalable scattering imaging.

a positive value represents a positive correlation, and 0 represents an irrelevant correlation. Since the optimization direction of the deep learning is optimized in the direction of loss value reduction, to obtain a positive reconstruction result, the NPCC is used for training [52]. The loss functions can be formulated as

$$\text{Loss} = \text{Loss}_{\text{NPCC}} + \text{Loss}_{\text{MSE}}, \quad (6)$$

$\text{Loss}_{\text{NPCC}}$

$$= \frac{-1 \times \sum_{x=1}^w \sum_{y=1}^b [i(x,y)] - \hat{i}][I(x,y)] - \hat{I}}{\sqrt{\sum_{x=1}^w \sum_{y=1}^b [i(x,y)] - \hat{i}}^2 \sqrt{\sum_{x=1}^w \sum_{y=1}^b [I(x,y)] - \hat{I}}^2}, \quad (7)$$

$$\text{Loss}_{\text{MSE}} = \text{Loss}_I = \sum_{x=1}^w \sum_{y=1}^b |\tilde{i}(x,y) - I(x,y)|^2, \quad (8)$$

where  $\hat{i}$  and  $\hat{I}$  are the mean value of the object ground truth  $I$  and the DNN output  $i$ , respectively, and  $\tilde{i}$  is a normalized image of  $i$ . The combination loss function has a good capability to reconstruct objects with different complexity and sparsity through different scattering media. To train the DNN, an Adam optimizer is selected as the strategy to update the weights in the training process. The DNN is performed on PyTorch 1.4.0 with a Titan RTX graphics unit and i9-9940X CPU under Ubuntu 16.04.

### 3. EXPERIMENTS AND RESULTS

#### A. Experimental Arrangement and Data Acquisition

The optical configuration is schematically illustrated in Fig. 3. A mounted LED (M625L4, Thorlabs, Newton, NJ, USA) and a filter (FL632.8-1, central wavelength:  $632.8 \pm 0.2$  nm, Thorlabs) are assembled as the light source, which can be employed as an approximate incoherent light source. A digital micromirror device (DMD) (pixel count:  $1024 \times 768$ , pixel size:

$13.68 \mu\text{m}$ ) is used to code and display the 8-bit objects. An industrial camera (acA1920-155um, Basler AG, Ahrensburg, Germany) is employed to obtain the patterns, which have lower data depth and relatively poor photo quality, instead of a scientific camera [8–13, 17–19, 52–54]. Thus, this method is more suitable for practical application. The ground glass is placed between the CMOS and DMD. The distance between the object and the diffuser is 30 cm, and the distance between the diffuser and the CMOS is 8 cm. The diameter of the iris behind the diffuser in the configuration is 8 mm, and the diameter of the iris combined with CL is 11 mm.

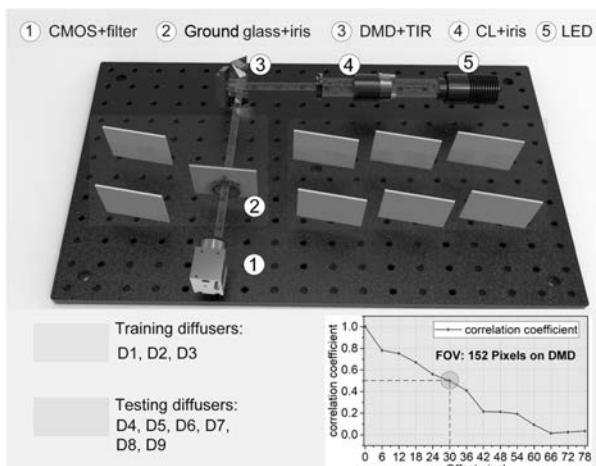
To obtain the speckle patterns in different scattering scenes, nine different ground glasses are used as the diffusers in the experiments, including six 220 grit diffusers (D1–D6), one 120 grit diffuser (D7), one 600 grit diffuser (D8) produced by Thorlabs, and one 220 grit diffuser (D9) produced by Edmund, like the configuration in Section 2.B. We choose one ground glass (D1) or the first three pieces of ground glasses (D1, D2, and D3) as the training diffusers, and the remaining ground glasses as the test diffusers. The objects are mainly selected from the MINIST database and FEI face databases. The character objects are selected randomly from the MINIST dataset to form the different complexity of single-character and double-character objects. For collecting the experimental data, 600 single characters, 600 double characters, and 400 human faces are used as objects hidden behind each diffuser. The first 500 characters are used as the seen objects and the remaining characters are used as the unseen objects. Similarly, the first 360 human faces are used as seen objects and the remaining faces are used as unseen objects. The autocorrelation pre-processing for speckle patterns is the first step for our method. As for the processing of the speckle patterns, we take the  $512 \times 512$  camera pixels from the center pattern to calculate the autocorrelation and crop the center to  $256 \times 256$  pixels autocorrelation pattern as the input autocorrelation image. All the objects, speckle patterns, and autocorrelation images are in grayscale in this experiment.

According to different training data and testing data, different groups are used to characterize the generalization capability of the physics-informed DL method, respectively. All of the testing data are captured from unknown diffusers for emphasizing the generalization. The data can be roughly divided into four groups.

**Group 1:** The objects are the single characters within the OME. The training data can be divided into two types: training with one diffuser (D1) or three diffusers (D1–D3) with seen objects (the first 500 characters). The testing data can also be divided into two types: the seen objects and the unseen objects (the last 100 characters) with testing diffusers (D4–D9).

**Group 2:** The objects are the double characters within the OME. The data arrangement is similar to Group 1, except for the complexity of objects.

**Group 3:** The objects are the human faces within the OME. The training data can also be divided into two types: training with one diffuser (D1) or three diffusers (D1–D3) with seen objects (the first 360 faces). The testing data can also be divided into two types: the seen objects and the unseen objects (the last 40 faces) with testing diffusers (D4–D9).



**Fig. 3.** Experimental setup for the scalable imaging. Different diffusers are employed to obtain speckle patterns with different scattering scenes. The OME range of this system is also measured by calculating the cross-correlation coefficient [21]. See Appendix B for details.

**Group 4:** The objects are the single characters extending the FOV to 1.2 times. The data arrangement is also similar to Group 1, except for the size and distribution of objects.

### B. Scalable Imaging with Different Diffusers

After collecting and classifying the data into different types, the proposed method is used for training and testing. The objects within OME (i.e., the first three groups), are tested first, and the experimental variables are the numbers of training diffusers and the category of the objects with different complexity and sparsity. To prove the good generalization capability and robustness of the physics-informed learning method, the subjective evaluation with reconstruction and the statistical average of objective evaluation results are provided in this section. Before the quantitative evaluating, the output images of the model have first been normalized. The imaging results shown in this paper are randomly selected from testing data, and the mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) are employed to quantitatively evaluate the generalization results and are presented in Table 1 with different groups. The results with Group 1 are presented with multiple examples in Fig. 4. Even with one training diffuser (D1), reliable generalization imaging results with unknown diffusers can be obtained, and the average PSNR is up to 23.41 dB. As shown in the reconstruction results of 0 in Fig. 4, it can be clarified that the seen objects testing results or unseen objects testing results with three training diffusers have higher fidelity than one training diffuser. Therefore, by increasing the quantity of training diffusers (D1-D3), the method can obtain higher accuracy and better generalization capability in unknown scattering scenes, and the average PSNR can reach over 40 dB with the training data of three diffusers.

To further verify the effectiveness of this method, more complex double-character objects and human faces are selected successively in this experiment. In addition to the single-character objects commonly used in traditional scattering scenes, double-character objects and FEI face database are selected in this paper for scheme verification, and the structures are more complex and suitable for actual application scenes. As shown in Fig. 5, double-character objects formed by combining single characters randomly can also be restored accurately through unknown diffusers. In the same way, by increasing

Ground truth	Training diffusers									Testing diffusers									Generalization results												
	Seen objects testing									Unseen objects testing																					
D1	D4	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D6	0	1	2	3	4	5	5	7	8	9	0	1	2	3	4	5	7	8	9	0	1	2	3	4	5	7	8	9	0	1
	D7	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D8	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	9	7	8	9	0	1	2	3	4	5	9	7	8	9
	D4	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D6	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D7	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D8	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9

**Fig. 4.** Testing results for generalization reconstruction of Group 1. Scale bars: 264.24 μm.

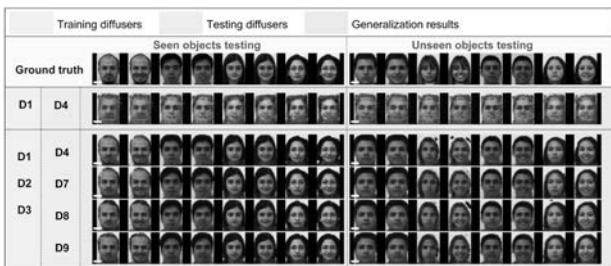
Ground truth	Training diffusers									Testing diffusers									Generalization results												
	Seen objects testing									Unseen objects testing																					
D1	D4	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D6	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D7	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D8	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D4	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D5	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D6	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D7	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D8	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9
	D9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9

**Fig. 5.** Testing results for generalization reconstruction of Group 2. Scale bars: 264.24 μm.

the quantity of training diffusers, the generalization capability can also be significantly improved. Compared to the character objects, the FEI faces database is more complex and difficult for the learning method. As shown in the second row in Fig. 6, the results of D4 are chosen as examples to show that the DL framework cannot reconstruct the human faces efficiently with

**Table 1.** Quantitative Evaluation Results of the Objects within OME

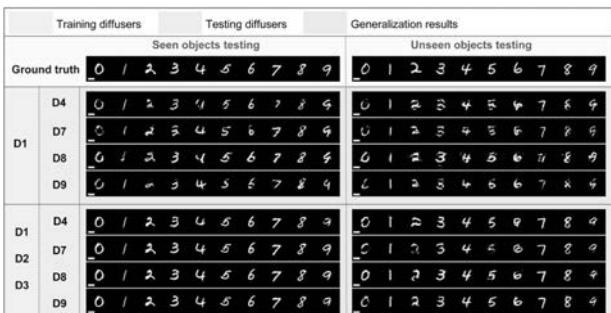
Type	Training Set	Testing Set	MAE	SSIM	PSNR (dB)
<b>Group 1</b>	D1	Seen objects	0.0190	0.8477	23.41
	D1	Unseen objects	0.0288	0.8248	19.86
	D1, D2, D3	Seen objects	0.0033	0.9744	40.56
	D1, D2, D3	Unseen objects	0.0271	0.8088	20.58
<b>Group 2</b>	D1	Seen objects	0.0257	0.8608	20.75
	D1	Unseen objects	0.0548	0.6929	16.56
	D1, D2, D3	Seen objects	0.0072	0.9190	38.60
	D1, D2, D3	Unseen objects	0.0455	0.7678	16.76
<b>Group 3</b>	D1	Seen objects	0.1166	0.4672	15.59
	D1	Unseen objects	0.1234	0.5351	14.98
	D1, D2, D3	Seen objects	0.0323	0.8523	23.82
	D1, D2, D3	Unseen objects	0.0534	0.7410	20.67



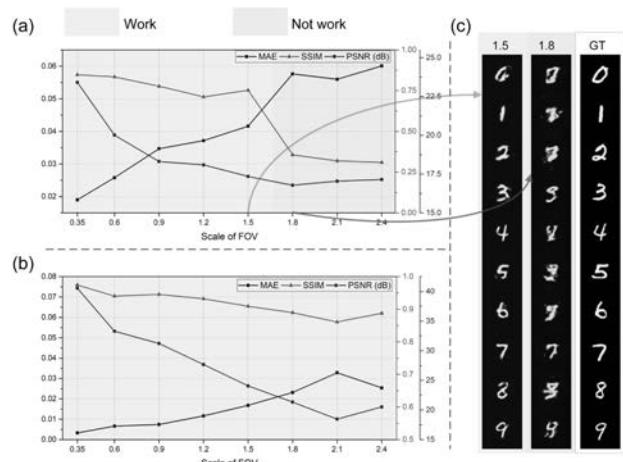
**Fig. 6.** Testing results for generalization reconstruction of Group 3. Scale bars: 264.24  $\mu\text{m}$ .

only one training diffuser. Once the quantity of training diffusers is improved to three diffusers, the generalization results are more accurate and reliable. In addition, one person has two photos with different micro-expressions in the FEI face database. From the seen objects testing results, with three training diffusers, the facial features and details of human faces can be clearly reconstructed. The scalable imaging of human faces has a high reduction degree, and the reconstructed faces with slight micro-expressions can also be identified and accurately distinguished.

When the object's scale exceeds OME, the object information is also contained by the speckle pattern and can be described with the speckle-correlation theory as Eq. (4). If the scale of the object exceeds the FOV of OME, it is hard for the traditional speckle-correlation methods to recover the object via a single speckle pattern. Using the powerful data-mining capabilities of CNN, the proposed method can extend the scope of OME, and have generalization capability for the large-scale objects shown in Fig. 7. As the quantity of training diffusers increases, the generalization effect can also be improved accordingly. From the quantitative evaluation results



**Fig. 7.** Testing results for generalization reconstruction of Group 4. Scale bars: 820.8  $\mu\text{m}$ .



**Fig. 8.** Generalization results for a single-character object with different scales and the scale of FOV is defined as the FOV/OME times. (a), (b) Results with different amounts of training diffusers, which are trained with one diffuser and three diffusers, respectively. (c) Reconstruction results with different scales and corresponding ground truth (GT).

in Table 2, the difficulty of scalable imaging exceeding OME is higher than Group 1, and the conclusion of the generalization capability is similar to the results within OME.

The FOV of scalable imaging beyond OME is affected by different configurations, such as the amount of the training diffusers, the complexity and sparsity of objects, and the camera hardware parameters. The generalization results with different scales are presented in Fig. 8, and the generalization conclusion with FOV is that the imaging effects and indicators are decreasing when the FOV is extending. As shown in Fig. 8(c), when the scale of the FOV extends to 1.8 times, the reconstructed characters are blurry and hard to distinguish. As shown in Fig. 8(b), with the quantity of training diffusers increasing, a similar conclusion can be drawn that the objective indicators of the generalization results are obviously improving.

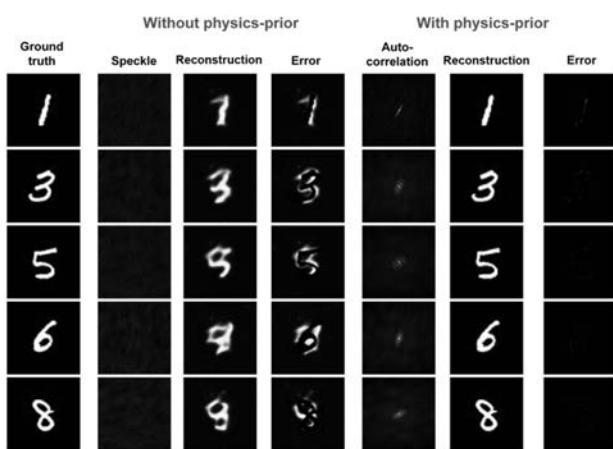
## 4. ANALYSIS

### A. Comparison to Traditional DL Strategy

To demonstrate the necessity of the physics-informed pre-processing step in imaging through unknown diffusers, the comparison images recovered by the end-to-end DL method without physics prior are presented in Fig. 9. As the aforementioned conclusions from Fig. 1, the speckle characteristics with different diffusers have a big difference. The unreliable imaging results are obtained from speckle patterns directly without

**Table 2. Quantitative Evaluation Results of Objects Extending the FOV 1.2 Times**

Type	Training Set	Testing Set	MAE	SSIM	PSNR (dB)
<b>Group 4</b>	D1	Seen objects	0.0359	0.8439	17.53
	D1	Unseen objects	0.0432	0.6968	17.12
	D1, D2, D3	Seen objects	0.0085	0.9431	30.55
	D1, D2, D3	Unseen objects	0.0309	0.8706	18.90

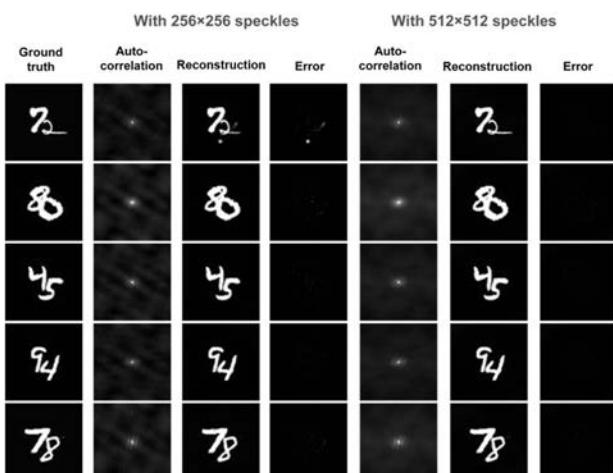


**Fig. 9.** Comparison results without or with this pre-processing step for imaging through an unknown diffuser. Three ground glasses are selected as the training diffusers and another diffuser for testing.

physics prior, and the reconstruction indicators are 0.2111 in SSIM and 15.54 dB in PSNR. Although there are a few objects that can be distinguished, such as the digits “1” and “3”, it is still hard for DNN to learn the speckle-correlation pre-processing step inside hidden layers automatically without an effective physics prior.

### B. Performance with Different Number of Speckles

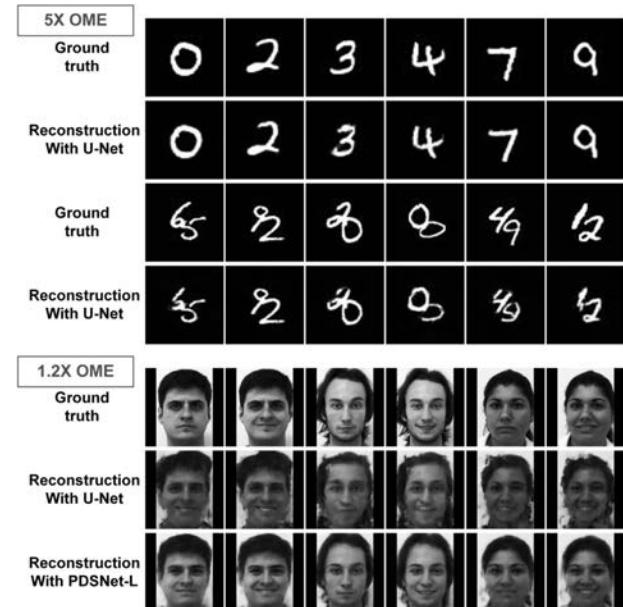
As shown in Fig. 9, the end-to-end DL method without physics prior can hardly work with raw speckle patterns for imaging through unknown diffusers. Thus, using the speckle-correlation prior is an effective step for the generalization of the imaging through unknown scattering media, and traditional speckle theory in speckle-correlation methods is also suitable to enhance the performance of physics-informed learning. Using relatively more speckles is an efficient way to reduce the statistical noise of the autocorrelation through scattering media [21]. The comparison results are shown in Fig. 10, using more speckles can also improve the performance of



**Fig. 10.** Results with different number of speckles via the physics-informed learning method. Three ground glasses are selected as the training diffusers and another diffuser for testing.

**Table 3.** Objective Indicators with Different Number of Speckles via the Physics-Informed Learning Method

Number of Speckles	MAE	SSIM	PSNR (dB)
256 × 256	0.0054	0.9342	38.38
512 × 512	0.0037	0.9539	44.53



**Fig. 11.** Generalization results of imaging exceeding OME range with different complexity objects.

the physics-informed learning method, and the objective indicators are also presented in Table 3.

### C. Performance in Exceeding FOV

Using the existing optical system, the imaging capability beyond the FOV of OME is also tested up to five times. As shown in Fig. 11, when the size of the digits is 760 × 760 pixels on DMD for a 5× OME range, the hidden objects can also be accurately predicted, and the objective indicators are presented in Table 4. We can obtain reliable results through an unknown scattering medium with three training diffusers, and the recovery difficulty increases with more complex objects. As for more complex objects (e.g., the FEI face dataset), the physics-informed learning method with a traditional U-Net cannot obtain reliable results. However, we can improve the generalization ability by using the PDSNet-L [55] for the

**Table 4.** Objective Indicators Corresponding to Fig. 11

Objects	Network	MAE	SSIM	PSNR (dB)
5× OME single characters	U-Net	0.0232	0.8741	21.70
5× OME double characters	U-Net	0.0361	0.8360	18.96
1.2× OME faces	U-Net	0.0673	0.6972	18.40
1.2× OME faces	PDSNet-L	0.0567	0.7050	21.16

neural network post-processing, which has a better reconstruction capability. Furthermore, we can further improve the generalization ability by using a better camera (a scientific CMOS or an electron-multiplying CCD) or gathering more training data.

## 5. DISCUSSION

According to the experimental results shown in Section 3, we have three key points.

(i) A physics-informed DL framework is proposed for scalable imaging through different scattering scenes in which the diffusers are previously untrained. The objects hidden behind the unknown diffusers are not limited to simple sparse characters, and more complex objects (e.g., human faces) can be reconstructed with high accuracy. The physics-informed learning method can also extend the FOV of conventional speckle-correlation methods.

(ii) The DL framework has a reliable generalization capability in imaging through unknown thin scattering media using only one training diffuser for the sparse object. With the number of training diffusers increasing, the generalization capability of the methods is further improved. The proposed method can still reconstruct the overall structure and local details for human faces, even the slight micro-expressions can be clearly distinguished. However, the DL models are prone to preferentially fit the category of the training dataset, which limits the generalization capability of the physics-informed learning method with unknown category objects.

(iii) Benefiting from the great capability in data-mining and mapping of DNN, reliable generalization results can also be obtained through unknown diffusers with the extended FOV. Meanwhile, the FOV of the physics-informed learning method is also relevant to several factors, such as the number of training diffusers and the complexity of the hidden objects.

## 6. CONCLUSION

In this paper, a physics-informed learning method is introduced to imaging through diffusers. Specifically, an explicit framework is established to efficiently solve the generalization problems in different scattering scenes by combining the physics theories and DL methods. This is a new approach to solve scalable imaging with deep learning, which can reconstruct complex objects through different scattering media, and provide an expanded FOV for the real imaging scenes. In the future, more complex scenes and objects can be considered, which can be applied to volumetric multiple scattering, such as biological imaging and astronomical imaging.

## APPENDIX A: THE FORMULA DERIVATION TO EXCEED THE OME RANGE

When the object size exceeds the range of OME, the object can be divided into multiple objects  $O_i$  within OME scope, and the PSFs produced from the different points become uncorrelated mutually. The autocorrelation of PSF can be approximately expressed as

$$\text{PSF}_i \star \text{PSF}_j \approx \begin{cases} \delta_{ij}, & i = j \\ 0, & i \neq j \end{cases} \quad (\text{A1})$$

We assume that the distance between objects is beyond the single OME range. Taking the autocorrelation of the camera image and using the convolution theorem yields [68]

$$\begin{aligned} I \star I &= \left( \sum_{i=1}^n O_i * \text{PSF}_i \right) \star \left( \sum_{i=1}^n O_i * \text{PSF}_i \right) = O_1 \\ &\star O_1 + C_1 + O_2 \star O_2 + C_2 + O_3 \\ &\star O_3 + C_3 + 2(O_1 \star O_2) * (\text{PSF}_1 \star \text{PSF}_2) + 2(O_2 \star O_3) \\ &* (\text{PSF}_2 \star \text{PSF}_3) + 2(O_1 \star O_3) * (\text{PSF}_1 \star \text{PSF}_3) + \dots \\ &= \sum_{i=1}^n (O_i \star O_i + C_i) = \sum_{i=1}^n (O_i \star O_i) + C'. \end{aligned} \quad (\text{A2})$$

Thus, the autocorrelation distribution of a speckle pattern exceeding OME can be defined as

$$I \star I = \sum_{i=1}^n (O_i \star O_i) + C'. \quad (\text{A3})$$

## APPENDIX B: OME RANGE CALIBRATION DETAILS

To calibrate the range of the shift-invariant, the distance from the object to the diffuser is changed to 15 cm and the image distance is maintained as 8 cm. A ground glass (DG100X100-220-N-BK7, Thorlabs) is used as the diffuser and placed between the object and the CMOS. A series of speckle patterns are collected while the horizontal displacement of the point object on the object surface is achieved. The cross-correlation coefficient between the speckle patterns and the PSF of the system is calculated. A threshold value of 0.5 is chosen as the cross-correlation coefficient to determine the range of OME [68,69]. We define  $\delta p$  as the offset pixel number of the image plane, which is 30 pixels, as shown in Fig. 3. The OME range of the system can be calculated by  $2 \times p \times \delta p / \beta$  [68], and  $\beta$  is the system magnification and  $p$  is the pixel size of the camera, which equals 5.86  $\mu\text{m}$ . It can be figured that the half-width at half maximum (HWHM) is 30 pixels. Because the distance from the object to the diffuser of the speckle collection system is 30 cm, the HWHM of the speckle collection system is 60 pixels. Then, the full width at half maximum of the speckle collection system is 120 pixels. Thus, the OME range of our speckle collection system is 152  $\times$  152 pixels on DMD.

**Funding.** National Natural Science Foundation of China (62031018, 61971227); Jiangsu Provincial Key Research and Development Program (BE2018126); Postgraduate Research & Practice Innovation Program of Jiangsu Province (KYCX20\_0264).

**Acknowledgment.** The authors thank Qianying Cui, Yingjie Shi, Chenyin Zhou, Kaixuan Bai, and Mengzhang Liu for technical support and experimental discussions.

**Disclosures.** The authors declare no conflicts of interest.

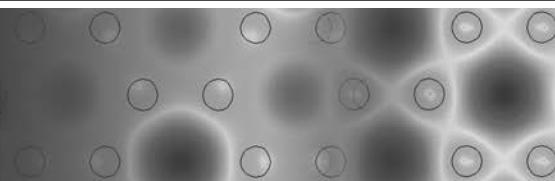
<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

1. J. W. Goodman, *Speckle Phenomena in Optics: Theory and Applications* (Roberts and Company, 2007).
2. M. C. Roggemann, B. M. Welsh, and B. R. Hunt, *Imaging Through Turbulence* (CRC Press, 1996).
3. R. K. Tyson, *Principles of Adaptive Optics* (CRC Press, 2015).
4. E. J. McCartney, *Optics of the Atmosphere: Scattering by Molecules and Particles* (Wiley, 1976).
5. V. Ntziachristos, "Going deeper than microscopy: the optical imaging frontier in biology," *Nat. Methods* **7**, 603–614 (2010).
6. S. Yoon, M. Kim, M. Jang, Y. Choi, W. Choi, S. Kang, and W. Choi, "Deep optical imaging within complex scattering media," *Nat. Rev. Phys.* **2**, 141–158 (2020).
7. L. V. Wang and H.-I. Wu, *Biomedical Optics: Principles and Imaging* (Wiley, 2012).
8. A. P. Mosk, A. Lagendijk, G. Lerosey, and M. Fink, "Controlling waves in space and time for imaging and focusing in complex media," *Nat. Photonics* **6**, 283–292 (2012).
9. I. M. Vellekoop and A. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**, 2309–2311 (2007).
10. S. Rotter and S. Gigan, "Light fields in complex media: mesoscopic scattering meets wave control," *Rev. Mod. Phys.* **89**, 015005 (2017).
11. K. Wang, W. Sun, C. T. Richie, B. K. Harvey, E. Betzig, and N. Ji, "Direct wavefront sensing for high-resolution *in vivo* imaging in scattering tissue," *Nat Commun* **6**, 7276 (2015).
12. S. Popoff, G. Lerosey, R. Carminati, M. Fink, A. Boccara, and S. Gigan, "Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media," *Phys. Rev. Lett.* **104**, 100601 (2010).
13. A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques," *Opt. Express* **23**, 11898–11911 (2015).
14. E. Tajahuerce, V. Durán, P. Clemente, E. Irles, F. Soldevila, P. Andrés, and J. Lancis, "Image transmission through dynamic scattering media by single-pixel photodetection," *Opt. Express* **22**, 16945–16955 (2014).
15. Y.-K. Xu, W.-T. Liu, E.-F. Zhang, Q. Li, H.-Y. Dai, and P.-X. Chen, "Is ghost imaging intrinsically more powerful against scattering?" *Opt. Express* **23**, 32993–33000 (2015).
16. Q. Fu, Y. Bai, X. Huang, S. Nan, P. Xie, and X. Fu, "Positive influence of the scattering medium on reflective ghost imaging," *Photon. Res.* **7**, 1468–1472 (2019).
17. D. Lu, M. Liao, W. He, Z. Cai, and X. Peng, "Imaging dynamic objects hidden behind scattering medium by retrieving the point spread function," *Proc. SPIE* **10834**, 1083428 (2018).
18. H. He, X. Xie, Y. Liu, H. Liang, and J. Zhou, "Exploiting the point spread function for optical imaging through a scattering medium based on deconvolution method," *J. Innov. Opt. Health Sci.* **12**, 1930005 (2019).
19. X. Xu, X. Xie, A. Thendiyamal, H. Zhuang, J. Xie, Y. Liu, J. Zhou, and A. P. Mosk, "Imaging of objects through a thin scattering layer using a spectrally and spatially separated reference," *Opt. Express* **26**, 15073–15083 (2018).
20. J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature* **491**, 232–234 (2012).
21. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**, 784–790 (2014).
22. A. Porat, E. R. Andrensen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, "Widefield lensless imaging through a fiber bundle via speckle correlations," *Opt. Express* **24**, 16835–16855 (2016).
23. J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
24. J. Chang and G. Wetzstein, "Single-shot speckle correlation fluorescence microscopy in thick scattering tissue with image reconstruction priors," *J. Biophoton.* **11**, e201700224 (2018).
25. P. Schniter and S. Rangan, "Compressive phase retrieval via generalized approximate message passing," *IEEE Trans. Signal Process.* **63**, 1043–1055 (2014).
26. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
27. I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning* (MIT, 2016), Vol. 1.
28. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
29. Y. Rivenson, Y. Zhang, H. Günaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**, 17141 (2018).
30. Y. Wu, Y. Rivenson, Y. Zhang, Z. Wei, H. Günaydin, X. Lin, and A. Ozcan, "Extended depth-of-field in holographic imaging using deep-learning-based autofocusing and phase recovery," *Optica* **5**, 704–710 (2018).
31. Z. Ren, Z. Xu, and E. Y. Lam, "Learning-based nonparametric autofocusing for digital holography," *Optica* **5**, 337–344 (2018).
32. Z. Ren, Z. Xu, and E. Y. Lam, "End-to-end deep learning framework for digital holographic reconstruction," *Adv. Photon.* **1**, 016004 (2019).
33. T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for Fourier ptychography microscopy," *Opt. Express* **26**, 26470–26484 (2018).
34. A. Kappeler, S. Ghosh, J. Holloway, O. Cossairt, and A. Katsaggelos, "Ptychnet: CNN based Fourier ptychography," in *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), pp. 1712–1716.
35. S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow," *Biomed. Opt. Express* **9**, 3306–3319 (2018).
36. Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, "Illumination pattern design with deep learning for single-shot Fourier ptychographic microscopy," *Opt. Express* **27**, 644–656 (2019).
37. M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," *Sci. Rep.* **7**, 17865 (2017).
38. Y. He, G. Wang, G. Dong, S. Zhu, H. Chen, A. Zhang, and Z. Xu, "Ghost imaging based on deep learning," *Sci. Rep.* **8**, 6469 (2018).
39. H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydin, L. A. Bentolila, C. Kural, and A. Ozcan, "Deep learning enables cross-modality super-resolution in fluorescence microscopy," *Nat. Methods* **16**, 103–110 (2019).
40. E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-storm: super-resolution single-molecule microscopy by deep learning," *Optica* **5**, 458–464 (2018).
41. W. Ouyang, A. Aristov, M. Lelek, X. Hao, and C. Zimmer, "Deep learning massively accelerates super-resolution localization microscopy," *Nat. Biotechnol.* **36**, 460–468 (2018).
42. C. Ling, C. Zhang, M. Wang, F. Meng, L. Du, and X. Yuan, "Fast structured illumination microscopy via deep learning," *Photon. Res.* **8**, 1350–1359 (2020).
43. L. Fang, D. Cunefare, C. Wang, R. H. Guymer, S. Li, and S. Farsiu, "Automatic segmentation of nine retinal layer boundaries in OCT images of non-exudative AMD patients using deep learning and graph search," *Biomed. Opt. Express* **8**, 2732–2744 (2017).
44. L. Waller and L. Tian, "Computational imaging: Machine learning for 3D microscopy," *Nature* **523**, 416–417 (2015).
45. T. C. Nguyen, V. Bui, and G. Nehmetallah, "Computational optical tomography using 3-D deep convolutional neural networks," *Opt. Eng.* **57**, 041406 (2018).
46. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**, 243902 (2018).
47. C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 3291–3300.
48. S. Feng, Q. Chen, G. Gu, T. Tao, L. Zhang, Y. Hu, W. Yin, and C. Zuo, "Fringe pattern analysis using deep learning," *Adv. Photon.* **1**, 025001 (2019).
49. K. Wang, Y. Li, Q. Kemao, J. Di, and J. Zhao, "One-step robust deep learning phase unwrapping," *Opt. Express* **27**, 15100–15115 (2019).

50. H. Yu, X. Chen, Z. Zhang, C. Zuo, Y. Zhang, D. Zheng, and J. Han, "Dynamic 3-D measurement based on fringe-to-fringe transformation using deep learning," *Opt. Express* **28**, 9405–9418 (2020).
51. H. Yu, D. Zheng, J. Fu, Y. Zhang, C. Zuo, and J. Han, "Deep learning-based fringe modulation-enhancing method for accurate fringe projection profilometry," *Opt. Express* **28**, 21692–21703 (2020).
52. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**, 803–813 (2018).
53. M. Lyu, H. Wang, G. Li, S. Zheng, and G. Situ, "Learning-based lensless imaging through optically thick scattering media," *Adv. Photon.* **1**, 036002 (2019).
54. N. Borhani, E. Kakkava, C. Moser, and D. Psaltis, "Learning to see through multimode fibers," *Optica* **5**, 960–966 (2018).
55. E. Guo, S. Zhu, Y. Sun, L. Bai, C. Zuo, and J. Han, "Learning-based method to reconstruct complex targets through scattering medium beyond the memory effect," *Opt. Express* **28**, 2433–2446 (2020).
56. E. Guo, Y. Sun, S. Zhu, D. Zheng, C. Zuo, L. Bai, and J. Han, "Single-shot color object reconstruction through scattering medium based on neural network," *Opt. Lasers Eng.* **136**, 106310 (2020).
57. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–1190 (2018).
58. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**, 16032–16046 (2019).
59. M. Liao, S. Zheng, D. Lu, G. Situ, and X. Peng, "Real-time imaging through moving scattering layers via a two-step deep learning strategy," *Proc. SPIE* **11351**, 113510V (2020).
60. Y. Li, S. Cheng, Y. Xue, and L. Tian, "Displacement-agnostic coherent imaging through scatter with an interpretable deep neural network," *Opt. Express* **29**, 2244–2257 (2020).
61. K. Goda, B. Jalali, C. Lei, G. Situ, and P. Westbrook, "AI boosts photonics and vice versa," *APL Photon.* **5**, 070401 (2020).
62. S. Feng, C. Kane, P. A. Lee, and A. D. Stone, "Correlations and fluctuations of coherent wave transmission through disordered media," *Phys. Rev. Lett.* **61**, 834–837 (1988).
63. I. Freund, M. Rosenbluh, and S. Feng, "Memory effects in propagation of optical waves through disordered media," *Phys. Rev. Lett.* **61**, 2328–2331 (1988).
64. H. Liu, Z. Liu, M. Chen, S. Han, and L. V. Wang, "Physical picture of the optical memory effect," *Photon. Res.* **7**, 1323–1330 (2019).
65. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, 2015), pp. 234–241.
66. Y. LeCun, C. Cortes, and C. J. C. Burges, "The MNIST database of handwritten digits," <http://yann.lecun.com/exdb/mnist/>.
67. C. E. Thomaz, "FEI face database," <https://fei.edu.br/~cet/facedatabase.html>.
68. C. Guo, J. Liu, W. Li, T. Wu, L. Zhu, J. Wang, G. Wang, and X. Shao, "Imaging through scattering layers exceeding memory effect range by exploiting prior information," *Opt. Commun.* **434**, 203–208 (2019).
69. D. Tang, S. K. Sahoo, V. Tran, and C. Dang, "Single-shot large field of view imaging with scattering media by spatial demultiplexing," *Appl. Opt.* **57**, 7533–7538 (2018).

# PHOTONICS Research



## Incoherent imaging through highly nonstatic and optically thick turbid media based on neural network

**SHANSHAN ZHENG,<sup>1,2</sup> HAO WANG,<sup>1,2</sup> SHI DONG,<sup>1,2</sup> FEI WANG,<sup>1,2</sup> AND GUOHAI SITU<sup>1,2,3,4,\*</sup>**

<sup>1</sup>*Shanghai Institute of Optics and Fine Mechanics, Chinese Academy of Sciences, Shanghai 201800, China*

<sup>2</sup>*Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China*

<sup>3</sup>*Hangzhou Institute for Advanced Study, University of Chinese Academy of Sciences, Hangzhou 310024, China*

<sup>4</sup>*CAS Center for Excellence in Ultra-intense Laser Science, Shanghai 201800, China*

\*Corresponding author: ghsitu@siom.ac.cn

Received 1 December 2020; revised 18 January 2021; accepted 18 February 2021; posted 19 February 2021 (Doc. ID 416246); published 21 April 2021

**Imaging through nonstatic scattering media is one of the major challenges in optics, and encountered in imaging through dense fog, turbid water, and many other situations. Here, we propose a method to achieve single-shot incoherent imaging through highly nonstatic and optically thick turbid media by using an end-to-end deep neural network. In this study, we use fat emulsion suspensions in a glass tank as a turbid medium and an additional incoherent light to introduce strong interference noise. We calibrate that the optical thickness of the tank of turbid media is as high as 16, and the signal-to-interference ratio is as low as -17 dB. Experimental results show that the proposed learning-based approach can reconstruct the object image with high fidelity in this severe environment.** © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416246>

### 1. INTRODUCTION

Conventionally, an optical imaging system can be regarded simply as a one-to-one mapping system, as a spherical wavelet emitted from a local point on the object plane is converged by the imaging optic and forms a unique point (subject to diffraction) on the imaging plane. In this way, a clear image of the object can be formed. However, the presence of scattering media between the object and the imaging optic prohibits such a clear image from being formed because a significant part of the light propagating all the way from the object to the optic is scattered many times, producing a noise-like scattered pattern on the image plane. As the optical thickness of the scattering medium is larger than a certain value, the scattered light arriving at the camera becomes more intensive than the unscattered or ballistic one; further, the image is submerged by the noise-like scattered pattern. Thus, an intuitive method for image acquisition under such hazed conditions is to select the ballistic light by using gating techniques in the time [1], space [2], polarization [3], or even coherence [4] domains.

Actually, the propagation and transport of coherent waves in disordered media has been one of the central problems in many different disciplines of science and engineering, ranging from microwaves to electron waves. It has been shown that coherent scattering in a static scattering medium is linear and deterministic and can be described by a transmission matrix [5]. One of

the lines of recent studies then focuses on the ways of reversing the deterministic coherent scattering process. Physically, this can be done by generating a phase conjugated version of the scattered wave [6,7], so that it can trace the way back through the scattering medium in the first place and form a clear image. An alternative way is to measure the transmission matrix [8] or precompensate the phase distortion introduced by the scattering process by the technique of wavefront shaping [5]. When applying these methods in dynamic scattering media, the wavefront measurement and playback must be made within the de-correlation time [9]. Therefore, faster wavefront modulators, parallel processing, and more effective optimization algorithms are highly demanded in these cases [10,11].

Alternatively, one can utilize some intrinsic statistical properties of the scattered light instead of reversing the scattering process. For example, the short-range second-order correlation of a speckle pattern formed by the scattered light, or the memory effect [12,13], has been found to be particularly interesting in this case, as it allows computational reconstruction of an object image from the autocorrelation of the scattered light intensity by using a regular phase-retrieval algorithm [14,15]. However, the memory effect offers a field of view (FoV) that is inversely proportional to the effective thickness of the medium. Thus, this method is usually applicable for imaging through optically thin scattering media [14,15] or around

the corner [16]. We note that efforts have been made to enlarge the FoV by using multiplexing or scanning [17,18]. But these techniques have limited effect in the case when the scattering medium is not static, as the dynamic characteristics of the media and the spectral width of the light in many practical environments will further shrink the correlation length of the scattered light [15,19] and thereby the FoV of the imaging system.

An interesting and practically important question then naturally arises, i.e., if it is possible to reconstruct an image when the coherence of the scattered light is completely lost. At least two factors may lead to the loss of coherence: multiple scattering of the light during its propagation inside the scattering medium [19] and the coherence nature of the light itself.

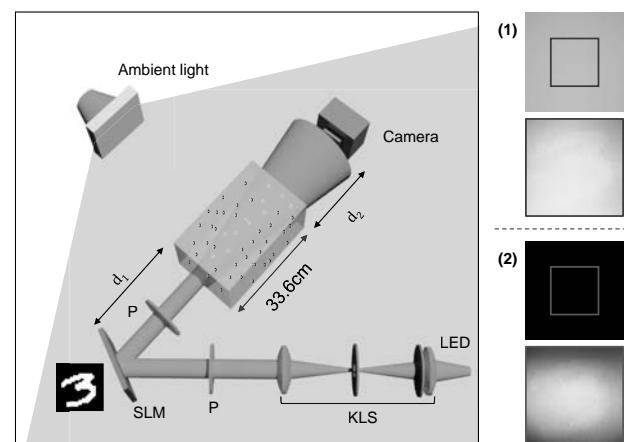
To answer this question, we attempt to demonstrate that it is still possible to reconstruct the image of an object hidden behind a nonstatic and optically thick diffusive medium and under the illumination of incoherent light. To mimic a realistic environment, we also switch on an ambient light so as to introduce a strong interference noise that is incoherently superimposed onto the scattered light from the object. The method we use to address this problem is deep neural networks.

Deep neural networks have shown great potential in solving many computational optical imaging problems [20], ranging from digital holography reconstruction [21,22], phase imaging [23,24], computational ghost or single-pixel imaging [25,26], to coherent imaging through optically thin [27,28] and even thick scattering media [29,30]. As almost all the deep neural networks for computational optical imaging were trained supervisedly (except, for example, the work reported in Ref. [24]), it is therefore required that each pair of labeled images in the training set should have exclusively unique features [31]. In the application of imaging through scattering media, when the illumination is coherent [27–30], what is recorded by the camera is a high-contrast speckle pattern, the morphology of which is explicitly object-dependent. However, when the objects are illuminated with an incoherent light, as in our case, the coherent effect will be smeared out, yielding milky patterns that are almost visually indistinguishable with respect to one another. Therefore, additional care has to be taken to identify unique features to distinguish each milky pattern.

This article is organized as follows. The problem of incoherent imaging through highly dynamic turbid media is formally described in Section 2. Measurements of the parameters that characterize the turbid medium are presented in Section 3. The experimental demonstration of the proposed method is presented in Section 4.

## 2. DESCRIPTION OF THE PROBLEM AND IMAGING ENVIRONMENT

Let us describe the problem with the help of Fig. 1. An incoherent light emitted from an LED (M530L3, Thorlabs) with a central wavelength of 530 nm was first collimated by using a Köhler illumination system and then shined on an object [which was displayed on an amplitude-only spatial light modulator (SLM; Pluto 6001, Holoeye)]. The light reflected from the SLM was then guided to a glass tank of fat emulsion suspension, which was prepared by mixing a tank of distilled water



**Fig. 1.** Incoherent scattering imaging experimental system. (1) and (2) are the captured scattered patterns (the raw data and corresponding partial contrast stretched map) with optical thickness of 8 and 16, respectively. Note that these data are recorded in two sets of experiments: (1) capture data by the camera directly; (2) capture data with two additional apertures placed before the camera. KLS, Köhler lighting system; P, polarizer; ambient light, generated by a high-power LED through a diffuse slate (the distance between the slate and the tank side was around 3.5 cm); camera, working with an imaging lens ( $f = 250$  mm, not shown in the figure).  $d_1 \approx 41$  cm,  $d_2 \approx 15$  cm. The 33.6 cm thick tank is equipped with fat emulsion diluent to simulate a dynamic scattering medium. Note that the scattered patterns shown in (2) look dimmed because a significant part of the large-angle scattered light has been blocked out.

and fat emulsion (Intralipid 20%, Fresenius Kabi). Multiple scattering events occur as the signal-carrying light propagated through the turbid medium and became highly diffusive when it left the tank (geometric length: 33.6 cm) from the back surface. An sCMOS camera (Zyla, Andor Technology Ltd.) with an  $f = 250$  mm imaging lens mounted on it was used to take the scattered light pattern at the back surface of the tank. According to the propagation trajectory, the signal scattered light included a small amount of weakly/snake light and a significant part of large-angle scattered light. The snake light contains a small amount of the object information and contributes to the signal on the imaging plane. The camera was adjusted to focus on the SLM plane before adding intralipid into the tank. We also introduced strong interference noise generated by a high-power LED (Solis-525c, Thorlabs) illuminated from one side of the tank. The central wavelength of the ambient LED is 525 nm. This beam of light was also scattered by the intralipid suspensions, and some of it was superimposed with the object scattered light incoherently, which can be treated as interference noise. Although the light was scattered mostly forwardly in the fat emulsion suspensions, the large angle and multiple-scattered ambient light was strong in comparison with the object scattered light. Indeed, the measured signal-to-interference ratio was  $SIR = -17$  dB in our experimental configuration.

The optical thickness of the tank of fat emulsion suspensions was about 16 (refer to Section 3 for the details of calibration). As the absorption of the intralipid is low, we can conclude that it is the scattering that mostly accounts for the attenuation

of the light. Then, the object's scattered light intensity  $\mathcal{S}\{I_o\}$  can be expressed as

$$\mathcal{S}\{I_o\} = \mathcal{S}_W\{I_o\} + \mathcal{S}_L\{I_o\}, \quad (1)$$

where  $\mathcal{S}_W\{I_o\}$  and  $\mathcal{S}_L\{I_o\}$  are the intensity of weak light and large-angle scattered light, respectively, without considering the ambient light.

In our experimental demonstration, we also take the ambient scattered light  $I_a$  into account. Thus, the final intensity pattern captured by the camera can be written as

$$I_s = \mathcal{S}_W\{I_o\} + \mathcal{S}_L\{I_o\} + I_a, \quad (2)$$

where the interference noise  $I_n = \mathcal{S}_L\{I_o\} + I_a$ , and  $\mathcal{S}_W$  can be approximately regarded as the signal component of the recorded image  $I_s$ . Thus, the signal-to-interference ratio SIR =  $10 \log_{10} \mathcal{S}_W\{I_o\}/I_n = -17$  dB.

Thus, the purpose is to reconstruct the object image  $I_o$  from  $I_s$ .

### 3. CHARACTERIZATION OF THE INTRALIPID SUSPENSIONS

Fat emulsion is similar to milk in many ways and mainly consists of soybean oil, water, glycerin, and eggphospholipid. Due to emulsification in the presence of phospholipids, the oil is suspended in water as small droplets with a monolayer lipid membrane. As a large amount of fat droplets are randomly distributed in the emulsion suspensions, a beam of light that propagates through it is distorted owing to multiple scattering, making it difficult to predict the propagation trajectory for each light ray.

Owing to the advantages such as low absorption, easily adjustable scattering coefficient by dilution, standardized particle size distribution, nontoxicity, and low cost [32], intralipid has been frequently used as an optical phantom that mimics a turbid medium, in particular, for biomedical applications. For example, it has been widely used in the calibration of clinical application systems [33,34] and as a diffusive reference standard for optical tissue modeling [35,36]. Many studies have been carried out to measure the scattering properties of intralipid suspensions [32,37]. However, when it is used as a scattering medium in imaging, the decorrelation time and optical thickness are quite important parameters related to the scattering, which are rarely measured or quantified. Thus, we need to calibrate these two parameters in this section.

#### A. Optical Thickness

The optical thickness (OT) is a parameter deduced from the Lambert–Beer law:

$$I = I_0 e^{-OT} = I_0 \cdot e^{-\mu \cdot c \cdot L}, \quad (3)$$

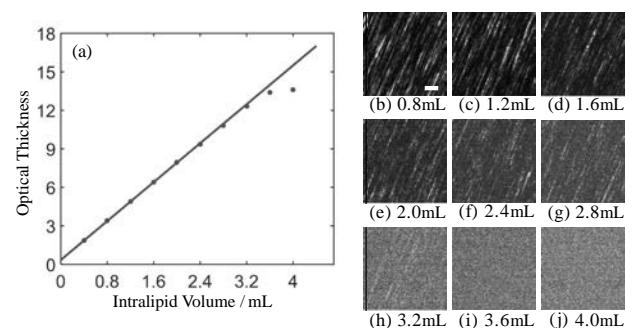
where  $I_0$  and  $I$  are the ballistic light intensities measured in front of and behind the tank of intralipid suspensions,  $\mu$  is the attenuation coefficient,  $c$  the concentration of intralipid dilution, and  $L$  the path length through the suspensions. The attenuation coefficient  $\mu$  is the sum of scattering coefficient  $\mu_s$  and the absorption coefficient  $\mu_a$ , i.e.,  $\mu = \mu_s + \mu_a$ .

Since we used the 530 nm LED (M530L3, Thorlabs) in our demonstrations, the optical thickness of the fat emulsion should be measured at or at least closed to this wavelength.

We thus used a semiconductor pumped green laser with a wavelength of 532 nm as the light source in the calibration process. An optical power meter (PM200, Thorlabs) with the probe diameter of 9.5 mm was used to measure  $I_0$  and  $I$ . Then, the optical thickness can be calculated according to Eq. (3). Note that the contributions to the absorption of fat emulsion are mainly from water and soybean oil, while the influence of glycerin and egg lipid can be ignored due to the low concentration. For the light at around 530 nm, the absorption coefficient  $\mu_a$  is about  $10^{-4} \text{ mm}^{-1}$  according to a previous study [32], so that its influence can be ignored; therefore,  $\mu \approx \mu_s$ .

In the measurement, we first filled in the glass tank with a volume of 11.46 L purified water ( $33.6 \text{ cm} \times 19.5 \text{ cm} \times 17.5 \text{ cm}$ ) and then gradually dropped 0–4 mL fat emulsion (intralipid 20%, Fresenius Kabi) into the water to make the intralipid suspensions with different density. Each time, about 0.4 mL intralipid was dropped into the tank of water. Then, we stirred the mixed fluid with a stick, let it stand for about 5 min, and then used a power meter placed on the beam path to measure the power of  $I_0$  and  $I$ . A large black cardboard with a 3 mm diameter aperture was used to select the ballistic light when measuring  $I$ . Then, the optical thickness can be calculated according to Eq. (3), and the result is plotted in Fig. 2(a). One can see that the experimental data linearly fit the theory very well. From the slope of the fitted line segment, we can obtain the scattering coefficient of the intralipid 20%,  $\mu_s = 132 \text{ mm}^{-1}$ , which is consistent with previous studies [32]. Deviation of the measurement data from theoretical prediction is observed when the density is high. This is because of the unmeasurable contribution of some amount of scattered light received by the power probe. Nonetheless, we can conclude that the optical thickness of the fat emulsion suspensions was about 16 when 4 mL intralipid was applied. This is the turbid medium that we want to look through.

The averaged diameter of the droplet is around  $0.3 \mu\text{m}$  [38]; at a scale comparable with the wavelength of the light, Mie scattering occurs when light passes through the fat emulsion suspension. During our experiment, we observed the Tyndall effect from the side of the tank, as shown later in this paper, which was taken by a mobile phone camera at the experiment site. Every droplet in the fat emulsion can be seen as a secondary source that re-emits a wavelet, the superposition of which forms a speckle pattern. However, we observed an interesting



**Fig. 2.** Optical thickness of intralipid suspensions with respect to its density. (b)–(j) Speckle patterns corresponding to different densities. Scale bar:  $200 \mu\text{m}$ .

phenomenon, i.e., the appearance of the speckle pattern changes with respect to the density of the intralipid suspensions. When the density is low, the speckle grain has a shape like a thin stripe [Fig. 2(b)–2(e)]. The reasons for this can be addressed as follows. First, the density of fat emulsion is low so that only a few scattering events occur to the light that propagates through it. That is why the contrast of the speckle in this case is very high compared with Figs. 2(i) and 2(j). Second, the droplets are moving downward owing to the gravity of the Earth during the course of data acquisition. Finally, the camera plane may be tilted with respect to the incoming light [39]. As the intralipid density increases, many more scattering events occur as the light propagates through the tank. As a result, the speckle pattern is formed by the interference of light waves that experienced various times of scattering. The speckle grain becomes more isotropic, but the contrast is apparently reduced (to 0.08).

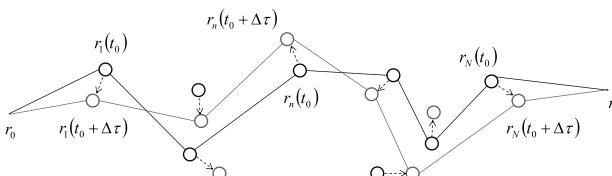
### B. Decorrelation Time

As noted, a large number of fat droplets are randomly distributed in emulsion suspension. Owing to gravity and Brownian motion in the surrounding liquid [40], these droplet scatterers are not static but moving randomly, resulting in fluctuations of light propagation trajectory in the turbid suspensions. For analyzing the influence of particle motion on the light scattering in this case, one usually attempts to measure the temporal correlation by using diffusing wave spectroscopy (DWS) [40–42]. In this paper, we propose an alternative method to characterize the temporal correlation property of the intralipid diluents with different concentrations in the 33.6 cm thick glass tank.

As illustrated in Fig. 3, when the scatterers move randomly during time interval  $\Delta\tau$ , the light propagation path, including multiple scattering events, will undergo unpredictable changes. It is obvious that the path fluctuation reflects the motion of the scatterer. Intuitively, we need to evaluate the temporal correlation between electric fields  $E(t_0 + \Delta\tau)$  and  $E^*(t_0)$  at point  $r$  (with a source placed at  $r_0$ ), which are closely related with the paired trajectories, as represented in Fig. 3. However, since traditional image sensors only record intensity, it is in principle not possible to directly detect the electric fields. According to the Siegert relation [43], the intensity correlation function can be associated with the electric field correlation function as

$$\frac{\langle I(t_0)I(t_0 + \Delta\tau) \rangle}{\langle I(t_0) \rangle \langle I(t_0 + \Delta\tau) \rangle} - 1 = \beta \left| \frac{\langle E(t_0 + \Delta\tau)E^*(t_0) \rangle}{\langle E(t_0 + \Delta\tau) \rangle \langle E^*(t_0) \rangle} \right|^2, \quad (4)$$

where  $\beta$  is an experimental factor between 0 and 1. Obviously, the temporal correlation property of the nonstatic medium can



**Fig. 3.** Multiple scattering trajectories in dynamic media. In this illustration, scatterers move from the black circle to the blue circle during time interval  $\Delta\tau$ , and  $r_i$  ( $i = 1, \dots, n, \dots, N$ ) represents the location where scattering event occurs.

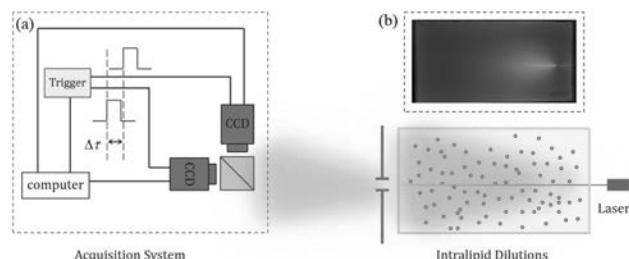
be characterized by using intensity-only measurements and calculate the intensity correlation between the speckle patterns taken at time  $t_0$  and at a later time  $t_0 + \Delta\tau$ .

The experimental system we designed to measure the temporal correlation of the speckle patterns is shown in Fig. 4. A semiconductor pumped laser emitted at 532 nm (MLL-S-532-A, Changchun New Industries Optoelectronics Tech. Co., Ltd.) was used as the incident light to illuminate the intralipid suspensions. The scattered light was divided into two arms by a beam splitter (BS) and collected, respectively, by two cameras (GS3-U3-23S6M-C, Point Grey). The dual cameras were controlled by an external trigger to achieve precise adjustment of the time lag of the two camera shutters. As shown in Fig. 4(a), the external control board Arduino generated two rectangular signals with a time interval  $\Delta\tau$  to trigger the two cameras, each of which acquired a speckle pattern,  $I(t_0)$  and  $I(t_0 + \Delta\tau)$ . This allows us to calculate the correlation coefficient between these two speckle patterns:

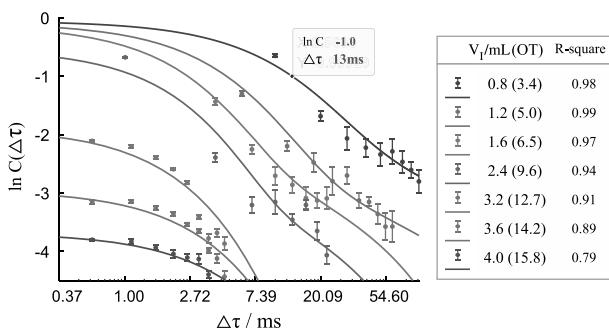
$$C(\Delta\tau) = \frac{\overline{\delta I(t_0) \cdot \delta I(t_0 + \Delta\tau)}}{\overline{\delta I(t_0)} \cdot \overline{\delta I(t_0 + \Delta\tau)}}, \quad (5)$$

where  $\delta I = I - \bar{I}$ , and  $\bar{I}$  is the mean of  $I$ . This function is basically identical with the temporal intensity correlation on the left-hand side of Eq. (4). By controlling the time delay  $\Delta\tau$  via the Arduino controller, it is therefore possible to calculate the curve of  $C(\Delta\tau)$  with respect to different  $\Delta\tau$ , manifesting the decorrelation behavior of the suspensions. In our experiment, the Arduino controller is set to trigger the two cameras with a minimum time lag of 100  $\mu$ s. In this way, we can measure the decorrelation behaviors of the intralipid suspensions with different concentrations. The experimental results are plotted in Fig. 5.

In order to determine the decorrelation time, we need to fit the data of  $C(\Delta\tau)$  to a theoretical model. Note that there are various contributions to the correlation function that are determined by the different sets of multiple scattering trajectories, which are classified in terms of the crossing number of two difusive paths [44]. Each contribution has discrepant time-dependence owing to different pairings of two trajectories obtained at time  $t_0$  and  $t_0 + \Delta\tau$ , respectively. One dominant contribution,  $C^{(1)}$ , the one without crossings, decreases exponentially with  $\Delta\tau$ . Further, it specifies the memory effect [12,13]. Another considerable contribution,  $C^{(2)}$ , has one crossing, which decreases algebraically [45]. Generally speaking, for one more path crossing, the corresponding



**Fig. 4.** Experiment setup. (a) Dual camera acquisition system. (b) Experimental site map of intralipid dilution: 11.47 L purified water (33.6 cm  $\times$  19.5 cm  $\times$  17.5 cm) and 2 mL intralipid 20%.



**Fig. 5.** Decorrelation curves for different concentrations of intralipid dilutions. The data points and the error bars represent the mean value and the standard error of the correlation coefficient calculated from 10 image pairs. The solid lines in different colors are the fitting results, and the corresponding intralipid volume  $V_I$  and optical thickness (OT) are shown in the legend. Here, the coefficient of determination ( $R$ -square) is used to describe the goodness of fit. Note that the horizontal axis is logarithmic scale.

contribution is reduced by a factor  $g$ , which is a large dimensionless number with a value at the order of  $10^2$  for a visible light in liquid suspensions [44,45]. As the contributions of other correlation terms with multiple crossings can be neglected, we only consider the effect of  $C^{(1)}$  and  $C^{(2)}$  here. Specifically, these two terms can be expressed as [19]

$$C^{(1)}(\Delta\tau) = \left[ \frac{L/L_s}{\sinh(L/L_s)} \right]^2, \quad (6)$$

$$C^{(2)}(\Delta\tau) = \frac{1}{g \sinh^2(L/L_s)} \left[ \frac{\sinh(2L/L_s)}{2L/L_s} - 1 \right], \quad (7)$$

where  $L$  is the geometric thickness of sample, and

$$L_s = \sqrt{D\tau_e} \cdot f(\Delta\tau), \quad (8)$$

where  $D$  is the diffusion coefficient that describes wave diffusion and is related to the elastic mean free path  $l_e$ ,  $\tau_e$  is the elastic collision time, and

$$f(\Delta\tau) = \left[ \frac{e^{-\Delta\tau/(2\tau_b)}}{1 - e^{-\Delta\tau/(2\tau_b)}} \right]^{1/2}, \quad (9)$$

where  $\tau_b$  describes the Brownian motion of the scatterers. One should note that  $\tau_b$  is the characteristic time for a scatterer to move a distance in the order of wavelength  $\lambda$ .

We fit the experimental data with the weighted sum of  $C^{(1)}$  and  $C^{(2)}$  defined in Eqs. (6) and (7), respectively, with the fitting parameters  $a$ ,  $b$ ,  $m$ , and  $n$  as

$$C(\Delta\tau) = a \left\{ \frac{m/f(n \cdot \Delta\tau)}{\sinh[m/f(n \cdot \Delta\tau)]} \right\}^2 + b \frac{1}{\{\sinh[m/f(n \cdot \Delta\tau)]\}^2} \left\{ \frac{\sinh[2m/f(n \cdot \Delta\tau)]}{2m/f(n \cdot \Delta\tau)} - 1 \right\}. \quad (10)$$

Apparently, the parameters  $a$  and  $b$  represent the related contributions of  $C^{(1)}$  and  $C^{(2)}$  to  $C$ . And  $m$  is a parameter related to  $L/\sqrt{D\tau_e}$ , while  $n$  is associated with  $\tau_b$ . The fitting results are plotted as solid lines in Fig. 5. From these fitted curves, one can

easily determine the decorrelation time  $\tau_d$ , which is defined as the time where the correlation coefficient drops to  $1/e$  of its maximum. We found that  $\tau_d$  for the suspensions with 0.8, 1.2, 1.6, and 2.4 mL intralipid 20% are 13, 4.8, 2.6, and 1.2 ms, respectively. Unfortunately, owing to the light being scattered too many times when we applied 3.6 mL intralipid or more in the tank, the decorrelation time drops dramatically, and the cameras are not fast enough to capture this process. We can conclude that the decorrelation time in this cases must be less than 100  $\mu$ s, which is the shutter speed limit of our cameras. Therefore, we assume that the decorrelation time of the turbid suspensions is around several to several tens of microseconds, corresponding to an optical depth of around 16.

## 4. EXPERIMENTAL DEMONSTRATIONS

### A. Experimental Data Acquisition

For the demonstration of the proposed method, we developed an experimental system, schematically shown in Fig. 1, to acquire all the data. In the first experiment, the optical thickness of the fat emulsion suspensions was 8 (2 mL intralipid 20%). In the second set of experiments, the optical thickness of the suspensions was about 16 (4 mL intralipid 20%), and decorrelation time was at the order of microseconds.

Here, we propose an end-to-end deep neural net to reconstruct object images from the acquired data. For convenience, we used 10,050 images altogether from the MNIST data set [46] as the training and test data. We first resized the original MNIST images to  $512 \times 512$  pixels and displayed them sequentially onto the central pixels of the  $1024 \times 1080$  pixel SLM. An incoherent LED light was used to illuminate the system. Reflected from the SLM, the incoherent light was then projected onto the front surface of the tank by using a delayed optic. As the light transmitted from the front surface all the way through the turbid suspensions to the back surface of the tank, it encountered many scattering events, as discussed in Section 3. The scattered patterns at the back surface of the tank were then captured by the sCMOS camera with an imaging optic. Note that there was also strong interference noise produced by an ambient light, with SIR = -17 dB, as described by Eq. (2). Only the central  $512 \times 512$  out of the  $1024 \times 1024$  pixels of the raw image taken by the camera were used in the reconstruction process. Our previous study [29] shows that this does not affect the quality of the reconstructed image. Owing to the limited frame rate of the SLM and the sCMOS camera, it took about 3.5 h to capture all 10,050 scattered patterns. Then, we partitioned them exclusively into two groups, one with 9900 patterns as the training set after pairing up with their corresponding ground true images (those displayed on the SLM), and the other as the test set. Typical visible features of the scattered patterns are shown in Fig. 1.

### B. Proposed Learning-Based Method for Image Reconstruction

The objective, then, is to reconstruct the object  $I_o$  from the corresponding recorded scattered pattern  $I_s$ . For the supervised training of a neural network, we need to construct a training set by pairing up every MNIST image in the training set together with the corresponding scattering pattern  $I_s$ ,

i.e.,  $S_{\text{training}} = \{(I_s^{(1)}, I_o^{(1)}), (I_s^{(2)}, I_o^{(2)}), \dots, (I_s^{(N)}, I_o^{(N)})\}$ . Then, we can converse the problem to learning the probability distribution  $P(I_o|I_s)$  [47] or the mapping relation  $R_{\text{learn}}: I_s \rightarrow I_o$  from the training set. Once it is learned, an object  $I_o$  in the test set should be predicted from the corresponding scattered pattern  $I_s$  [29]. The objective function then can be defined as

$$\mathcal{R}_{\text{learn}} = \arg \min_{\mathcal{R}_\theta, \theta \in \Theta} \sum_{n=1}^N \mathcal{L}(I_o^{(n)}, \mathcal{R}_\theta\{I_s^{(n)}\}) + g(\theta), \quad (11)$$

where the set  $\Theta$  contains handcraft parameters that specify the network structure and weighting parameters that are automatically adjusted during the training process [48],  $\mathcal{L}(\cdot)$  is the loss function that evaluates the error between label  $I_o^n$  and predicted output  $\mathcal{R}_\theta\{I_s^n\}$ , and  $g(\theta)$  is a regularizer defined on the parameters with the aim of avoiding overfitting [20].

We adopted the convolutional neural network (CNN) [46] to solve the problem described by Eq. (11). Refer to Appendix A for more details about the structure of the neural net we designed.

To train the neural net, we must specify the loss function in Eq. (11). In this work, we adopted the mean square error (MSE) as the loss function to evaluate the predicted network outputs and ground truth images:

$$\text{MSE} = \min \frac{1}{WHN_1} \sum_{n=1}^{N_1} \sum_{(u,v)} [I_p^{(n)}(u,v) - I_o^{(n)}(u,v)]^2, \quad (12)$$

where  $I_p^{(n)}(u,v)$  is defined as the reconstructed image from the  $n$ th scattered pattern  $I_s^{(n)}(u,v)$ ,  $I_o^{(n)}(u,v)$  is the corresponding ground-truth resized MNIST image,  $W$  and  $H$  are the width and height of the reconstructed image, and  $N_1$  is the batch size, which was set to 5 in this work. We used the stochastic gradient descent method [49] to evaluate the MSE and Adam [50] to adjust the weights. The training step number was set as 30,000; in practice, however, we were able to use less training steps.

### C. Experimental Results

The experimental results are shown in Figs. 6(b) and 6(c). In comparison with the corresponding ground truth images shown in Fig. 6(a), the MNIST images are reconstructed successfully. We use the root of mean square error (RMSE) and structural similarity index (SSIM) [51] to evaluate reconstructed images quantitatively. These functions are defined as

$$\text{RMSE} = \left\{ \frac{1}{WH} \sum_{(u,v)} [I_p(u,v) - I_s(u,v)]^2 \right\}^{\frac{1}{2}}, \quad (13)$$



**Fig. 6.** Experimental results. (a) Ground truths, and the reconstructed images in the case that the optical thickness equals (b) 8 and (c) 16, respectively.

$$\text{SSIM} = \frac{(2\mu_{I_p} \mu_{I_s} + c_1)(2\sigma_{I_p I_s} + c_2)}{(\mu_{I_p}^2 + \mu_{I_s}^2 + c_1)(\sigma_{I_p}^2 + \sigma_{I_s}^2 + c_2)}, \quad (14)$$

where  $\mu_I$  and  $\sigma_I^2$  are the mean and variance of the image  $I$ ,  $\sigma_{I_p I_s}$  is the covariance between  $I_p$  and  $I_s$ , and  $c_1$  and  $c_2$  are regularization parameters. Basically, the RMSE measures the averaged difference between the reconstructed and corresponding ground truth images, whereas the SSIM evaluates the structure similarity between the two images. We calculate the RMSE and SSIM of all the 150 test images. The averaged values are 5.25 and 0.99, respectively.

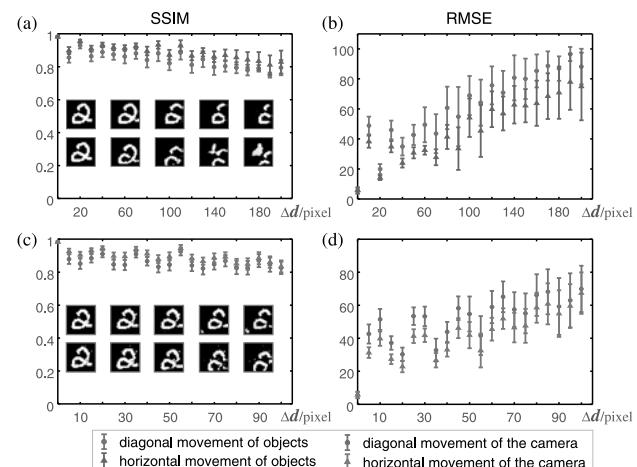
As mentioned above, we also performed a second set of experiments, in which the optical thickness of the turbid suspension is 16. In this case, we must block out a significant part of the large-angle scattered light by placing two apertures between the back surface of the tank and the camera. The captured scattered patterns then look dimmed, as shown in panel (2), Fig. 1. The corresponding reconstructed images from the proposed neural net are plotted in Fig. 6(c). With respect to the ground truths, the RMSE was about 4.66 and the SSIM was around 0.98, at the same level as those shown in Fig. 6(b).

### D. Robustness Analysis

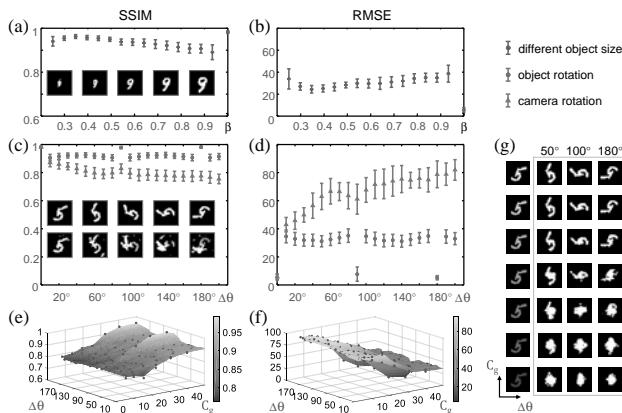
In this subsection, we concisely analyze the robustness of the proposed technique. The optical thickness of the turbid suspension is 16 in all the analysis here.

#### 1. Robustness against the Movement of Object/Camera

The first situation we considered is how the change of the object/camera position in the test with respect to that in the training affects the reconstructed image without retraining the neural net. The experimental results are shown in Figs. 7 and 8. In Fig. 7, we consider the separated effect of the change of position of the object and the camera, both along the diagonal and horizontal directions. One can see that, in all the cases, the images are reconstructed with the SSIM value better than 0.8. The proposed method seems to be more robust to the change of the object position owing to the scattering of light from the object. The RMSE values do not reflect the robustness



**Fig. 7.** Robustness against the position change of the object/camera.  $\Delta d$  is the displacement of the object/camera (in pixel). The data points and the error bars represent the mean values and the standard deviations of the SSIM/RMSE of 10 reconstructed images (digits '0–9').



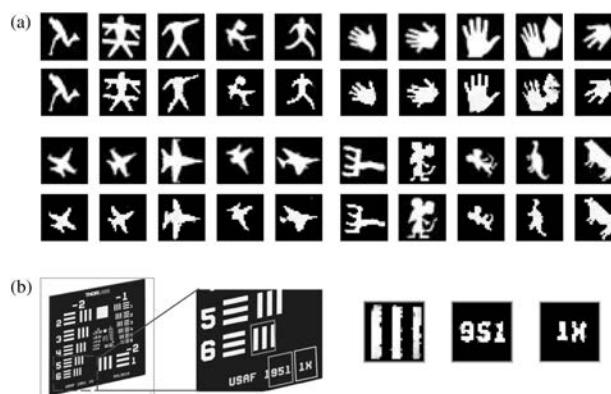
**Fig. 8.** Robustness against the scaling and rotation of the object/camera.  $\beta$  is the scaling factor of the image size,  $\Delta\theta$  the rotation angle, and  $C_g$  the image contrast gradient. The data points and the error bars in (a)–(d) represent the mean values and the standard deviations of the SSIM/RMSE of 10 reconstructed images (digits ‘0–9’). (e) and (f) SSIM/RMSE of digit ‘5’ with respect to  $\Delta\theta$  and  $C_g$ . (g) Visualized reconstructed digits.

in these cases because of the shifting of the reconstructed image, as shown in the insets of Figs. 7(a) and 7(c). In addition, artifacts arise when the object is shifted too far away, as can be seen in the insets in Fig. 7(a).

The robustness to the scaling and rotation of the object/camera is shown in Fig. 8. In both cases, the image can be reconstructed well. Again, it is more robust to the change in the object (both scaling and rotation) owing to the reason mentioned above.

## 2. Generalization to Nondigit Objects

We also used the trained neural net to reconstruct nondigit objects from the corresponding scattered patterns. The results are shown in Fig. 9. Two types of objects were used: virtual objects, which were displayed on the SLM [Fig. 9(a)], and a physical USAF target [Fig. 9(b)]. The experimental results



**Fig. 9.** Reconstruction of nondigit objects with the neural network trained by using digits. (a) First and third rows are the ground truths; second and fourth are the corresponding reconstructed images. (b) Reconstructed USAF target and the highlight of some of its portions.



**Fig. 10.** Experimental results with natural scene object. (a) Scattered patterns. (b) Corresponding ground truth. (c) Reconstructed results.

suggest that the proposed technique can be used to reconstruct objects that can be far from those that were used to train the neural net.

## 3. Reconstruction of Nature Images

However, the neural network has to be retrained if the object is more complex such as nature scenes that are in gray scale and have complex structures. However, it is still possible to reconstruct the image with a good fidelity. For demonstration, we tried the proposed method on a more complex set of objects. We took the STL-10 data set [52] in our experiment. Specifically, we used 19,000 images to retrain the network model and other 1000 images for testing. Some of the typical results are shown in Fig. 10. The averaged RMSE and SSIM values are 36.14 and 0.63, respectively.

## 5. CONCLUSION

In conclusion, we have demonstrated an end-to-end deep-learning-based incoherent imaging method through optically thick and dynamic scattering media under a condition of strong interference noise (SIR = -17 dB). The proposed method allows direct reconstruction of an object image from the captured scattered pattern with high quality (with an averaged RMSE around 5 and averaged SSIM about 0.63). We have quantitatively measured the optical thickness and decorrelation time of the turbid suspension [which was made by mixing fat emulsion (intralipid 20%) with purified water], with respect to the intralipid concentrations, and calibrated the optical thickness of the tank of intralipid suspensions in our experiments, which is about 16 in the severest case. Although the results were obtained in the lab environment, it is expected that the proposed method can be used in a wider scope of other scattering environments.

## APPENDIX A: CNN STRUCTURE

The proposed neural network structure is based on the one we published previously in Ref. [22] and shown Fig. 11. As can be seen, the neural network has a typical U-net structure [53] but with independent branches, which are designed to learn the training set features at different scales.

In our experiments, the central  $512 \times 512$  pixels of each captured scattered pattern  $I_s$  have to be preprocessed before input into the neural network. These images were first processed by three sets of convolution blocks and a max-pooling layer, resizing the image size to  $64 \times 64$  pixels. Then, the network was divided into four independent paths, each of which equips with a max-pooling layer to implement one, two, four, and

eight times downsampling of the incoming images, respectively, as indicated by the downward arrows in Fig. 11(a). For each data flow path, after the pooling layer, images were put to four identical residual blocks and an appropriate number of upsampling blocks, which was designed to make image size as  $64 \times 64$  pixels. Next, these independent paths were concatenated into an intact image, which contains 192 feature maps. It was then followed by a series of convolution blocks and three upsampling blocks to resize the output image to  $512 \times 512$  pixels. Finally, a convolution block was used to reduce the output channels to 1, so that the output results were gray-level images. The final output of our neural network was the reconstructed images  $I_o$  when the network was well-trained.

Now, we present a detailed description of the proposed neural network. There are mainly three types of functional blocks in the neural network: the convolution block, the residual block, and the upsampling block, as shown in Fig. 11(b). Compared with the neural network in Ref. [22], a batch normalization layer [54] is added into the convolution and residual blocks, so as to accelerate learning speed and be less sensitive to initialization. Furthermore, it is also expected to act as a regularizer, as described in Eq. (11). In our implementation, the convolution block contains a convolutional layer, a batch normalization layer, and an activation function. The convolution layer is the core layer to build a convolutional neural network and achieves feature extraction at different scales. The activation function used here is the rectified linear units. As shown in Fig. 11(a), a max-pooling layer is periodically inserted between

successive convolution layers, and its function is to gradually reduce the spatial dimension of the data and the number of internal weights required for image reconstruction. Here, max pool is with  $2 \times 2$  filters and stride 2, so the image size is reduced by a factor of 2. The residual block includes two consecutive groups consisting of a convolutional layer, a batch normalization layer, and an activation function. The shortcut connections between the input and the output enable us to optimize the neural network and improve accuracy by adding considerable depth. In addition, the upsampling block is composed of a transposed convolution layer and an activation function, which is to achieve image enlargement and convolution decoding. The stride of the transposed convolution layer is set as 2 to double the image size by zero-padding in between two adjacent pixels. Moreover, the detailed input channel numbers and output channel numbers of hidden layers in the neural network are represented as 1-16, 16-32, etc., as shown under the block symbols in Fig. 11(a). The size of convolutional kernels is presented as (5, 5) or (3, 3) below the channel numbers.

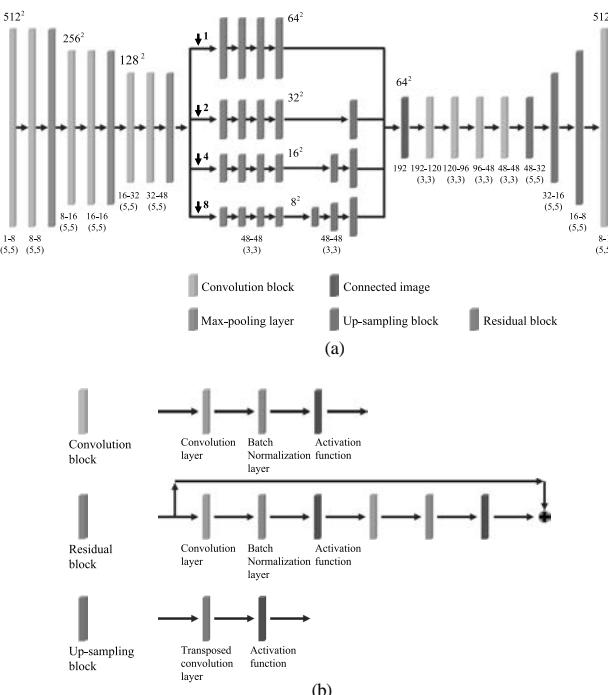
**Funding.** Chinesisch-Deutsche Zentrum für Wissenschaftsförderung (GZ1391); National Natural Science Foundation of China (61991452, 62061136005); Chinese Academy of Sciences Key Project (QYZDB-SSW-JSC002).

**Acknowledgment.** We thank the anonymous reviewers for their constructive comments.

**Disclosures.** The authors declare no conflicts of interest related to this paper.

## REFERENCES

1. L. Wang, P. P. Ho, C. Liu, G. Zhang, and R. R. Alfano, "Ballistic 2-D imaging through scattering walls using an ultrafast optical Kerr gate," *Science* **253**, 769–771 (1991).
2. Q. Z. Wang, X. Liang, L. Wang, P. P. Ho, and R. R. Alfano, "Fourier spatial filter acts as a temporal gate for light propagating through a turbid medium," *Opt. Lett.* **20**, 1498–1500 (1995).
3. S. G. Demos and R. R. Alfano, "Optical polarization imaging," *Appl. Opt.* **36**, 150–155 (1997).
4. E. N. Leith, C. Chen, H. Chen, Y. Chen, J. Lopez, P.-C. Sun, and D. Dilworth, "Imaging through scattering media using spatial incoherence techniques," *Opt. Lett.* **16**, 1820–1822 (1991).
5. I. M. Vellekoop, "Feedback-based wavefront shaping," *Opt. Express* **23**, 12189–12206 (2015).
6. Z. Yaqoob, D. Psaltis, M. S. Feld, and C. Yang, "Optical phase conjugation for turbidity suppression in biological samples," *Nat. Photonics* **2**, 110–115 (2008).
7. M. Cui and C. Yang, "Implementation of a digital optical phase conjugation system and its application to study the robustness of turbidity suppression by phase conjugation," *Opt. Express* **18**, 3444–3455 (2010).
8. S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, "Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media," *Phys. Rev. Lett.* **104**, 100601 (2010).
9. M. M. Qureshi, J. Brake, H.-J. Jeon, H. Ruan, Y. Liu, A. M. Safi, T. J. Eom, C. Yang, and E. Chung, "In vivo study of optical speckle decorrelation time across depths in the mouse brain," *Biomed. Opt. Express* **8**, 4855–4864 (2017).
10. D. Wang, E. H. Zhou, J. Brake, H. Ruan, M. Jang, and C. Yang, "Focusing through dynamic tissue with millisecond digital optical phase conjugation," *Optica* **2**, 728–735 (2015).



**Fig. 11.** Proposed neural network architecture. (a) Digits in the format  $m-n$  below each layer denote the number of input channels  $m$ , and the number of output channels  $n$ . (5, 5) and (3, 3) denote the size of the convolution kernel in pixel counts. (b) Detailed information of neural network structure.

11. L. Yan, M. Cheng, S. Yuecheng, S. Junhui, and L. V. Wang, "Focusing light inside dynamic scattering media with millisecond digital optical phase conjugation," *Optica* **4**, 280–288 (2017).
12. S. Feng, C. Kane, P. A. Lee, and A. D. Stone, "Correlations and fluctuations of coherent wave transmission through disordered media," *Phys. Rev. Lett.* **61**, 834–837 (1988).
13. I. Freund, M. Rosenbluh, and S. Feng, "Memory effects in propagation of optical waves through disordered media," *Phys. Rev. Lett.* **61**, 2328–2331 (1988).
14. J. Bertolotti, E. G. van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Non-invasive imaging through opaque scattering layers," *Nature* **491**, 232–234 (2012).
15. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**, 784–790 (2014).
16. C. A. Metzler, F. Heide, P. Rangarajan, M. M. Balaji, A. Viswanath, A. Veeraraghavan, and R. G. Baraniu, "Deep-inverse correlography: towards real-time high-resolution non-line-of-sight imaging," *Optica* **7**, 63–71 (2020).
17. T. Dongliang, S. S. Kumar, T. Vinh, and D. Cuong, "Single-shot large field of view imaging with scattering media by spatial demultiplexing," *Appl. Opt.* **57**, 7533–7538 (2018).
18. G. Li, W. Yang, H. Wang, and G. Situ, "Image transmission through scattering media using ptychographic iterative engine," *Appl. Sci.* **9**, 849 (2019).
19. E. Akkermans and G. Montambaux, *Mesoscopic Physics of Electrons and Photons* (Cambridge University, 2007).
20. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
21. Y. Rivenson, Y. Zhang, H. Güneydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," *Light Sci. Appl.* **7**, 17141 (2018).
22. H. Wang, M. Lyu, and G. Situ, "eHoloNet: a learning-based end-to-end approach for in-line digital holographic reconstruction," *Opt. Express* **26**, 22603–22614 (2018).
23. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117–1125 (2017).
24. F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light Sci. Appl.* **9**, 77 (2020).
25. M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," *Sci. Rep.* **7**, 17865 (2017).
26. F. Wang, H. Wang, H. Wang, G. Li, and G. Situ, "Learning from simulation: an end-to-end deep-learning approach for computational ghost imaging," *Opt. Express* **27**, 25560–25572 (2019).
27. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**, 803–813 (2018).
28. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–1190 (2018).
29. M. Lyu, H. Wang, G. Li, S. Zheng, and G. Situ, "Learning-based lensless imaging through optically thick scattering media," *Adv. Photon.* **1**, 036002 (2019).
30. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**, 16032–16046 (2019).
31. J. Gua, Z. Wangb, J. Kuenb, L. Mab, A. Shahroudyb, B. Shuaib, T. Liub, X. Wangb, and G. Wang, "Recent advances in convolutional neural networks," *Pattern Recogn.* **77**, 354–377 (2018).
32. R. Michels, F. Foschum, and A. Kienle, "Optical properties of fat emulsions," *Opt. Express* **16**, 5907–5925 (2008).
33. F. Martelli and G. Zaccanti, "Calibration of scattering and absorption properties of a liquid diffusive medium at NIR wavelengths. CW method," *Opt. Express* **15**, 486–500 (2007).
34. J. T. Allardice, A. M. Abulafi, D. G. Webb, and N. S. Williams, "Standardization of intralipid for light scattering in clinical photodynamic therapy," *Laser Med. Sci.* **7**, 461–465 (1992).
35. P. D. Ninni, F. Martelli, and G. Zaccanti, "Intralipid: towards a diffusive reference standard for optical tissue phantoms," *Phys. Med. Biol.* **56**, N21–N28 (2011).
36. B. W. Pogue and M. S. Patterson, "Review of tissue simulating phantoms for optical spectroscopy, imaging and dosimetry," *J. Biomed. Opt.* **11**, 041102 (2006).
37. M. Raju and S. N. Unni, "Concentration-dependent correlated scattering properties of intralipid 20% dilutions," *Appl. Opt.* **56**, 1157–1166 (2017).
38. M. L. Dong, K. G. Goyal, B. W. Worth, S. S. Makkar, W. Calhoun, L. M. Bali, and S. Bali, "Accurate *in situ* measurement of complex refractive index and particle size in intralipid emulsions," *J. Biomed. Opt.* **18**, 087003 (2013).
39. H. T. Yura, S. G. Hanson, R. S. Hansen, and B. Rose, "Three-dimensional speckle dynamics in paraxialoptical systems," *J. Opt. Soc. Am. A* **16**, 1402–1412 (1999).
40. G. Maret and P. E. Wolf, "Multiple light scattering from disordered media: the effect of Brownian motion of scatterers," *Z. Phys. B* **65**, 409–413 (1987).
41. M. J. Stephen, "Temporal fluctuations in wave propagation in random media," *Phys. Rev. B* **37**, 1–5 (1988).
42. D. J. Pine, D. A. Weitz, P. M. Chaikin, and E. Herbolzheimer, "Diffusing wave spectroscopy," *Phys. Rev. Lett.* **60**, 1134–1137 (1988).
43. P.-A. Lemieux and D. J. Durian, "Investigating non-Gaussian scattering processes by using *n*th-order intensity correlation functions," *J. Opt. Soc. Am. A* **16**, 1651–1664 (1999).
44. F. Scheffold and G. Maret, "Universal conductance fluctuations of light," *Phys. Rev. Lett.* **81**, 5800–5803 (1998).
45. F. Scheffold, W. Hartl, G. Maret, and E. Matijevic, "Observation of long-range correlations in temporal intensity fluctuations of light," *Phys. Rev. B* **56**, 10942–10952 (1997).
46. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE* **86**, 2278–2324 (1998).
47. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT, 2016).
48. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature* **323**, 533–536 (1986).
49. T. S. Ferguson, "An inconsistent maximum likelihood estimate," *J. Am. Stat. Assoc.* **77**, 831–834 (1982).
50. K. Diederik and J. Ba, "Adam: a method for stochastic optimization," arXiv:1412.6980 (2014).
51. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.* **13**, 600–612 (2004).
52. A. Coates, A. Y. Ng, and H. Lee, "An analysis of single-layer networks in unsupervised feature learning," *J. Mach. Learn. Res.* **15**, 215–223 (2011).
53. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)* (Springer, 2015), pp. 234–241.
54. S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," arXiv:1502.03167 (2015).

# PHOTONICS Research

## Realizing transmitted metasurface cloak by a tandem neural network

ZHENG ZHEN,<sup>1,2,3</sup> CHAO QIAN,<sup>1,2,3,5</sup> YUETIAN JIA,<sup>1,2,3</sup> ZHIXIANG FAN,<sup>1,2,3</sup> RAN HAO,<sup>4,6</sup> TONG CAI,<sup>1,2,3</sup> BIN ZHENG,<sup>1,2,3,7</sup> HONGSHENG CHEN,<sup>1,2,3</sup> AND ERPING LI<sup>1,2,3</sup>

<sup>1</sup>Interdisciplinary Center for Quantum Information, State Key Laboratory of Modern Optical Instrumentation, College of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China

<sup>2</sup>ZJU-Hangzhou Global Science and Technology Innovation Center, Key Laboratory of Advanced Micro/Nano Electronic Devices & Smart Systems of Zhejiang, Zhejiang University, Hangzhou 310027, China

<sup>3</sup>International Joint Innovation Center ZJU-UIUC Institute, Zhejiang University, Haining 314400, China

<sup>4</sup>College of Optical and Electronic Technology, China Jiliang University, Hangzhou 310018, China

<sup>5</sup>e-mail: chaoqianzju@zju.edu.cn

<sup>6</sup>e-mail: ran.hao@cjlu.edu.cn

<sup>7</sup>e-mail: zhengbin@zju.edu.cn

Received 6 January 2021; revised 4 March 2021; accepted 4 March 2021; posted 12 March 2021 (Doc. ID 418445); published 30 April 2021

Being invisible at will has been a long-standing dream for centuries, epitomized by numerous legends; humans have never stopped their exploration steps to realize this dream. Recent years have witnessed a breakthrough in this search due to the advent of transformation optics, metamaterials, and metasurfaces. However, the previous metasurface cloaks typically work in a reflection manner that relies on a high-reflection background, thus limiting the applications. Here, we propose an easy yet viable approach to realize the transmitted metasurface cloak, just composed of two planar metasurfaces to hide an object inside, such as a cat. To tackle the hard-to-converge issue caused by the nonuniqueness phenomenon, we deploy a tandem neural network (T-NN) to efficiently streamline the inverse design. Once pretrained, the T-NN can work for a customer-desired electromagnetic response in one single forward computation, saving a great amount of time. Our work opens a new avenue to realize a transparent invisibility cloak, and the tandem-NN can also inspire the inverse design of other metamaterials and photonics. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.418445>

### 1. INTRODUCTION

Perhaps no one feels unfamiliar with the possibility of an invisibility cloak, as this long-standing dream has been epitomized by numerous legends and novels for centuries. To realize this dream, humans have never stopped their exploration. Particularly, in the past decades, the advent of metamaterials and nanotechnology has ignited unprecedented enthusiasm in realizing manmade cloaks that brings a new twist to the conventional cloaking community [1–7]. A groundbreaking proposal of a transformation optics-based cloak renders an object invisible by bending the flow of light around it, suppressing the scattering to be exactly zero [2]. In theory, this method is perfect; however, in experiment, it is marred by the bulky material compositions with both anisotropy and inhomogeneity [3]. As the two-dimensional (2D) equivalence of metamaterials, metasurfaces recently have demonstrated their rich optical properties in providing abrupt phase shift, amplitude modulation, and polarization conversion of electromagnetic (EM) waves [8–15]. By covering a deliberately designed metasurface over the hidden object, the scattered fields can be

reconstructed as if an incident wave were to impinge onto a pure background without the hidden object [16–20]. The usage of metasurface greatly reduces the thickness and complexity of invisibility cloaks, but they typically work in a reflection manner, known as a carpet cloak, and rely on a high-reflection background as the reference. This dependence limits its applications.

The transmitted metasurface cloak can properly compensate for the imperfection of a reflected metasurface cloak and work without any background (free-standing invisibility cloak), i.e., creating an illusion of free space. Therefore, the EM wave should be either guided around the hidden object or compensated for by lossy/gain media, putting forward a higher requirement for the metasurface design. So far, there have been some proposals about transmitted metasurface cloaks; for example, Ref. [21] integrated transparent metasurfaces and zero-index materials together to realize a hybrid invisibility cloak in transmission geometry; Ref. [22] proposed a parity-time symmetric metasurface with balanced gain and loss to realize a cloak in a particular direction. These works greatly enrich the transmitted

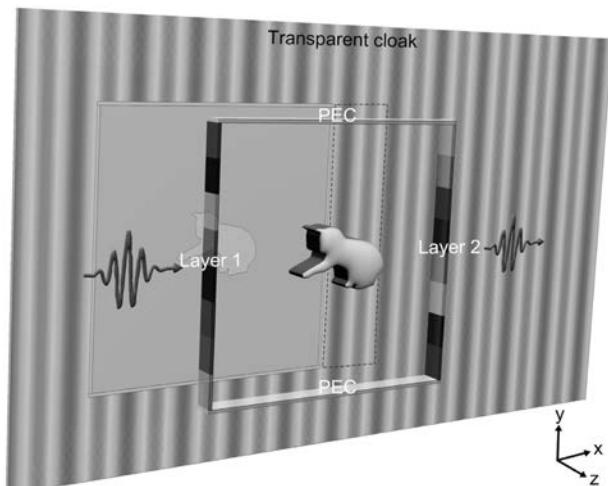
metasurface cloak, but they are complicated and cumbersome in practical implementation. Also, in designing these cloaks, the key step—metasurface design—conventionally relies on time-consuming EM numerical simulations, assisted by manual fine-tuning or optimization algorithms, such as genetic algorithms, to iteratively approach the demanded optical responses. Fundamentally, these stochastic algorithms work in a trial-and-error manner that are limited by their random search nature [23–30]. Given these factors, albeit challenging, a practical preferable transmitted metasurface cloak approach and an intelligent design algorithm have been highly sought after.

In this paper, an easy yet viable approach is proposed to realize the transmitted metasurface cloak, and a tandem neural network (T-NN) to efficiently streamline the inverse design process is introduced. The transmitted metasurfaces cloak is composed of two planar metasurfaces to hide an object inside, such as a cat. Compared with conventional methods, our transparent cloak does not involve any active component, which greatly simplifies the practical realization. The tandem-NN is used to address the hard-to-converge issue caused by the non-uniqueness phenomenon that widely exists in the EM inverse problem [27]. We collect the simulation data by the finite-element method analysis software COMSOL Multiphysics and train the network-NN with the open-source high-level deep learning application programming interface (API) Keras, with an accuracy of 86.6% and 86.5% for near and far fields, respectively. The pretrained T-NN can work for a customer-desired optical response, e.g., invisibility cloak, in one single forward computation. Our work provides a new avenue to realize a transparent cloak, distinct from conventional reflected metasurface cloaks. Also, the T-NN can inspire the inverse design of other metadevices [20,31,32].

## 2. RESULTS

**Architecture of the transmitted metasurfaces invisibility cloak.** As schematically shown in Fig. 1, our proposed transmitted metasurface cloak is easy yet viable, just consisting of two planar metasurfaces, i.e., metasurface 1 (left) and metasurface 2 (right). At the top and bottom sides, two perfect electric conductor (PEC) blocks are arranged to form a closed rectangular box to prevent EM waves from scattering in other directions. Inside the rectangular box, we assume there is a dielectric cat model with the relative permittivity of 3.5, which can also be generally replaced by other objects with different shapes and materials. For conceptual clarity, we assume the transparent cloak works for the transverse magnetic (TM) plane wave excited from the left side and at 8 GHz in two dimensions. Each metasurface is composed of eight subwavelength metasurface elements (with a width of 20 mm). Here, to accelerate the simulation and data collection process, we mimic the metasurface element by setting a transition boundary condition on the section line of its position. A metasurface element is enabled to provide a continuous local transmitted spectrum shift (0 to  $2\pi$ ) with unitary transmission. To characterize the cloaking performance, we consider both near-field distribution (the out-of-plane magnetic field inside the rectangle region enclosed by the green dashed line) and far-field radar cross section (RCS).

By modifying the metasurface properties, the near- and far-field

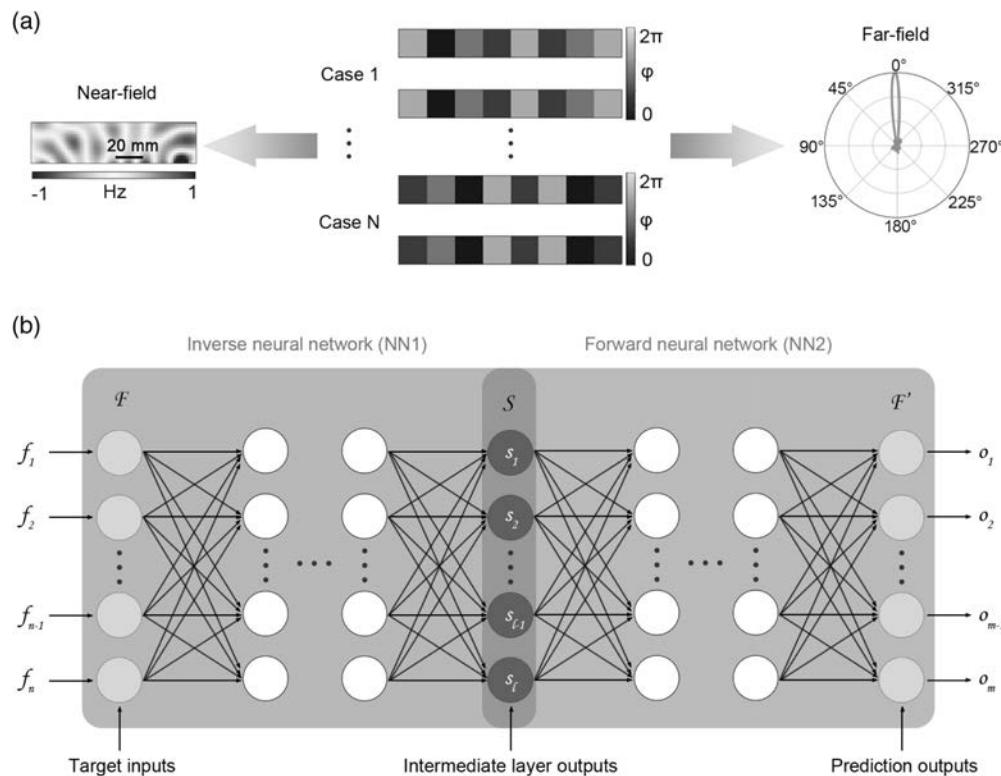


**Fig. 1.** Schematic of the transmitted metasurface cloak. The transmitted metasurface cloak consists of two planar metasurfaces, labelled as layer 1 and layer 2, to hide an object inside, such as a cat. Each metasurface is composed of eight subwavelength metasurface elements, each of which provides a local transmitted spectrum shift. To prevent EM waves from scattering in other directions, two PEC blocks are arranged to form a closed rectangular box. Here, we consider both near-field distribution (the out-of-plane magnetic field inside the rectangular region enclosed by the green dashed line) and far-field radar cross section to characterize the cloaking performance. Ideally, after passing through the two-layer metasurfaces, the forward and backward scattering of the incident wave shall be very small, as though the rectangular box were transparent.

signatures will change correspondingly. Ideally, after passing through the layered metasurfaces, the forward and backward scattering of the incident wave shall be very small, as though the rectangular box were transparent.

**Nonuniqueness issue and T-NN.** Evidently, our goal is to exploit and generalize the intricate inverse relationship between the near field/far field and the transmitted spectrum of metasurface using a deep neural network. However, there is a great challenge in training the deep neural network because of the existing nonuniqueness issue in the inverse design. To be specific, the same EM field response  $F$  can be generated by multiple different metasurface arrangements  $S$ . As showcased in Fig. 2(a), two different metasurface arrangements lead to exactly the same far field and near field. This nonunique  $F \rightarrow S$  will induce conflicting training samples, such as  $(F, S^A)$  and  $(F, S^B)$  in Fig. 2(a). Therefore, in the training process, the inverse neural network will encounter the serious hard-to-converge problem because the buildup data set will inevitably contain some cases having different output labels, but with the same input.

To overcome this nonunique issue, we deploy a T-NN, as shown in Fig. 2(b). The T-NN consists of two deep neural networks, i.e., an inverse neural network (termed NN1) and a forward neural network (termed NN2). For a customer-desired EM response  $F$ , the NN1 will generate a candidate of the metasurface arrangement  $S$ , and then feed  $S$  into the NN2 to predict the EM response  $F'$ . According to the uniqueness principle of electromagnetism, for the NN2, each kind of input  $S$



**Fig. 2.** Nonuniqueness issue addressed by a T-NN in the inverse design. (a) Different metasurface arrangements induce exactly the same near field and far field, called the nonuniqueness issue. This nonuniqueness issue will make the deep neural network difficult to converge. (b) Schematic of a T-NN, consisting of an inverse deep neural network (NN1) and a forward deep neural network (NN2). The NN1 has the input of near-/far-field response and the output of metasurface arrangement (nonuniqueness). In contrast, the NN2 has the input of metasurface arrangement and the output of near-/far-field response (uniqueness). In the training procedure of the T-NN, the NN2 is pretrained and fixed, and only the NN1 is updated to reduce the loss function, that is, the difference between the target field response  $F$  and the output  $F'$ . Therefore, the metasurface arrangement  $S$  can be extracted from the intermediate layer.

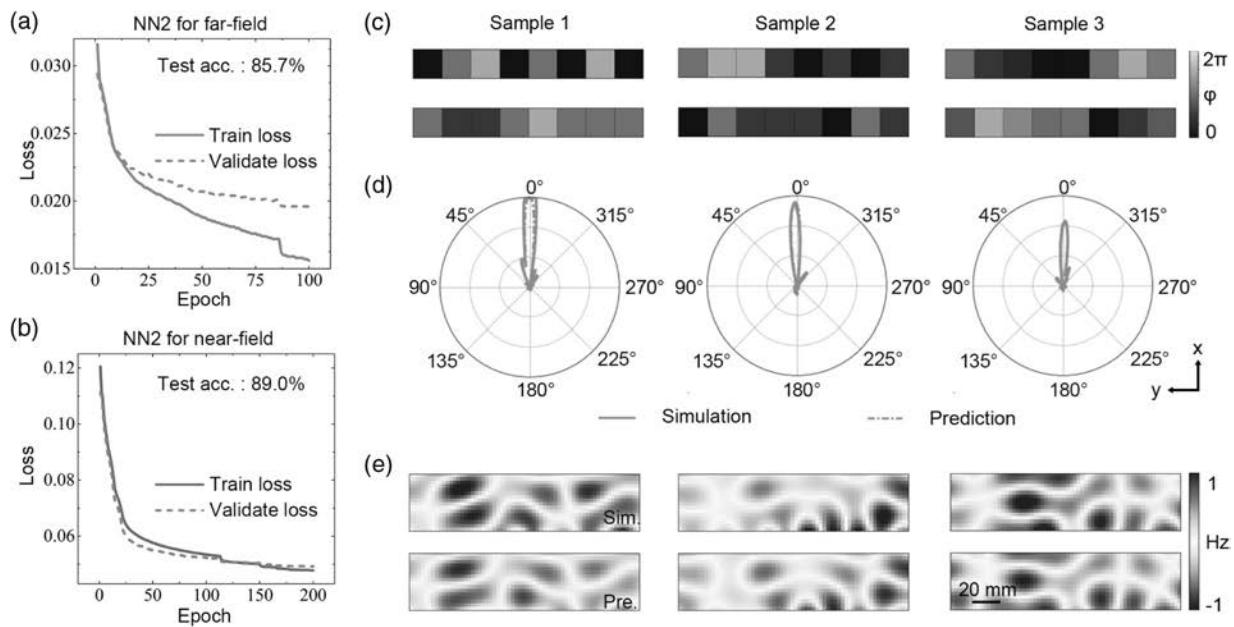
will only induce one unique and deterministic output, distinct from the NN1. The T-NN takes the difference between the input  $F$  and the NN2's output  $F'$  to calculate the loss function, and then iteratively updates the network. Once the T-NN is well trained, the predicted structural parameter  $S$  can be extracted from the intermediate layer output. In practice, we intend to obtain the metasurface arrangement with a customer-desired EM field response. Therefore, the T-NN can meet our practical demand and effectively address the hard-to-converge problem.

**Training and evaluation of the T-NN.** Before inversely designing the transparent cloak, we should consider how to build up the data set, train the T-NN, and evaluate them. To simplify the process, we only consider transmitted phase modulation while keeping the amplitudes uniform and unity (which can be accessible with current metasurface technology). Therefore, in our bilayer metasurface system, 16 metasurface elements correspond to 16 transmitted phases, each of which can be freely tuned from 0 to  $2\pi$ , with a minimum precision of  $0.01\pi$ . The random data of the whole phase distribution are generated step by step, i.e., we first construct some data seeds with a larger phase difference, and then use smaller phase difference to further generate data based on these seeds, so as to obtain better data distribution. The randomly generated structural

parameters are imported into the simulation software, generating a great number of far-field and near-field data. In total, a data set with 200,000 samples is built up, among which 70%, 20%, and 10% are used as the training, validation, and test set.

Before training the T-NN, the forward neural network, i.e., NN2, must be trained in advance, because of the one-to-one correspondence between the input  $S$  and the output  $F'$ . After optimization, the NN2 architecture is set to have seven hidden layers for the far field (five hidden layers for near field), and each layer has 1024 neurons. The input layer has 361 neurons (discretizing the elevation angle  $0 - 2\pi$  into 361 points) for the far field, and 1024 neurons (converting field distribution into  $16 \times 64$  image pixel matrix) for the near field.

Figure 3 plots the training results of the NN2, where the loss function gradually converges, and ultimately, the accuracy at the test set reaches 85.7% (89.0%) for the far field (near field). To intuitively show the training performance, we blindly take three samples from the test set [Fig. 3(c)], and predict the EM field responses by the pretrained NN2 [green line in Fig. 3(d)]. As a comparison, we also import the input metasurface arrangements into the commercial numerical software COMSOL to obtain the EM response [ground truth, red line in Fig. 3(d)]. Each group of far-field images in Fig. 3(d) [as well as in Fig. 4(c)] is processed by a uniform normalization to better



**Fig. 3.** Training results of the forward deep neural network (NN2). (a) Learning curve of the NN2 for the far field, with an accuracy of 85.7%; (b) learning curve of the NN2 for the near field, with an accuracy of 89.0%; (c) three metasurface arrangement samples, taken from the test set, to illustrate the performance of the NN2; (d) normalized RCS predicted by the NN2 and the simulated one obtained by importing the above three samples into the commercial numerical software COMSOL; (e) near-field distributions predicted by the NN2, and the simulated one obtained by numerical simulation.

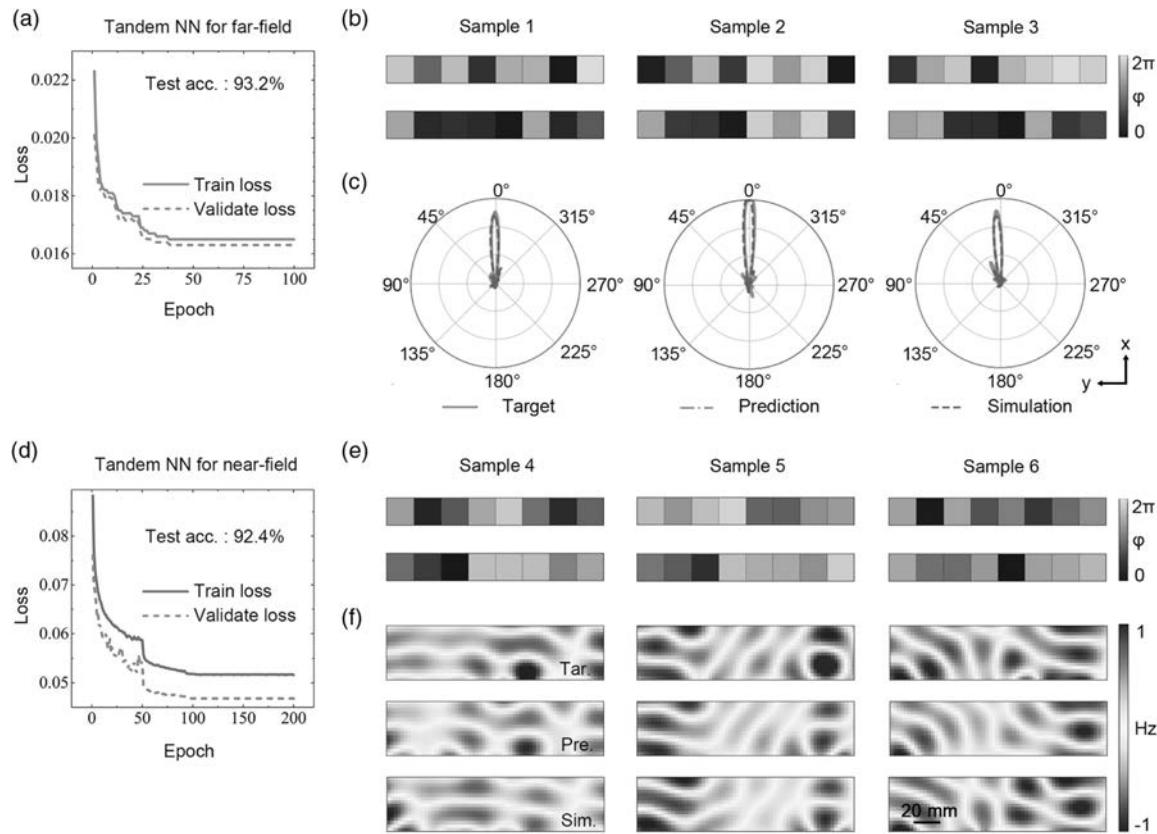
show the result match. As shown in Figs. 3(d) and 3(e), the predicted field patterns and the ground truths are highly consistent with each other, laying a foundation for the following training of the T-NN.

With the successful training of the NN2, we next connect it to the NN1 to form a complete T-NN [Fig. 2(b)]. After optimizing, the NN1 architecture is set to have eight hidden layers (seven 1024-neuron layers plus one 512-neuron layer) for the far field, and five 1024-neuron hidden layers for the near field. It is worth noting that, in this process, the NN2 is fixed and does not participate in the parameter updating. The EM field response  $F = [f_1, f_2, \dots, f_{361}]$  for the far field and  $F = [f_1, f_2, \dots, f_{1024}]$  for the near field is separately taken as the input to train the T-NN. As mentioned above, the training aims to minimize the loss function, defined as the mean absolute error (MAE) between  $F$  and  $F'$ . Figures 4(a)–4(c) are the training results for the far field, and Figs. 4(d)–4(f) are the training results for the near field. With the decrease of the loss function [Fig. 4(a)], the test accuracy achieves 93.2%. We utilize three samples to intuitively demonstrate the training result. Figure 4(c) exhibits three types of EM field response. Type 1 is the target RCS [red curve in Fig. 4(c)], Type 2 is the output EM field response generated by the T-NN [green curve in Fig. 4(c)], with the intermediate output metasurface arrangements  $S$  in Fig. 4(b). Type 3 is the numerical simulation of the  $S$  from the intermediate layer output [purple curve in Fig. 4(c)], which is presented as a comparison with Type 1 and Type 2. The three types of normalized RCS are drawn with the same coordinates as in Fig. 4(c), which match well with each other. For the near field [Fig. 4(d)], the loss function of the T-NN also reduces significantly, and the accuracy of

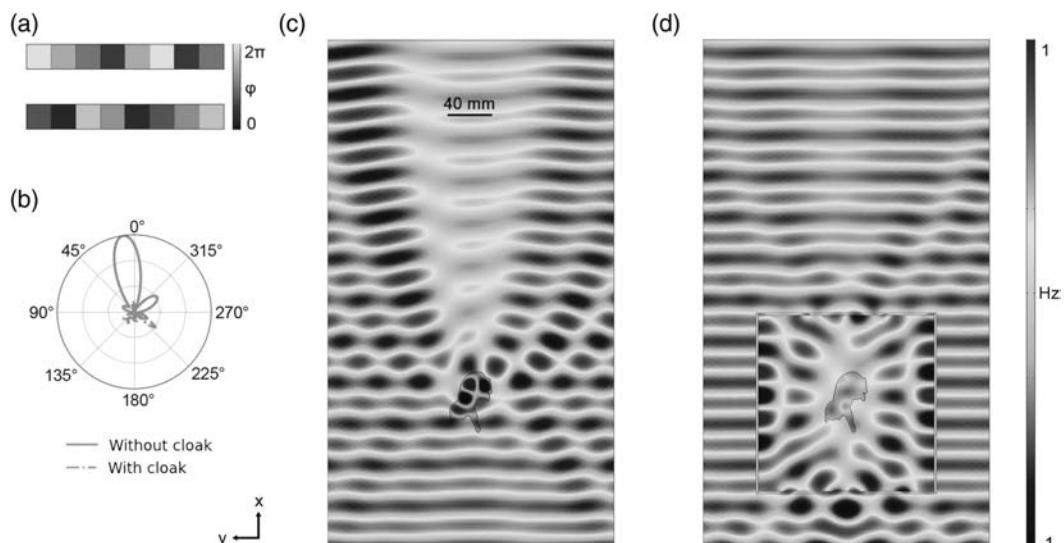
the test set reaches 92.4%. Similar to the far field, we consider three samples. The target magnetic fields are plotted in the upper panel of Fig. 4(f), the output magnetic fields from the T-NN are plotted in the middle panel of Fig. 4(f), and the simulated magnetic fields are plotted in the lower panel of Fig. 4(f). These samples validate that the tandem network has a strong generality, capable of working for a customer-desired EM response.

**Transparent invisibility cloak designed by the T-NN.** Based on the pretrained T-NN, we progress to utilize bilayer metasurfaces to realize a transparent invisibility cloak. We input the target far field and target near field into our T-NN to predict the metasurface arrangements. After a fast calculation, we get the desired transmitted phase  $\varphi$  arrangement shown in Fig. 5(a). Figure 5(b) clearly demonstrates that the RCS is significantly reduced in the forward direction when the invisibility cloak is present, in stark contrast to that without the cloak. More intuitively, regarding the near-field distribution, our bared cat-shaped object leads to a strong shadow in Fig. 5(c), while the scattered field is well reconstructed in Fig. 5(d). This way, our bilayer metasurface designed by the T-NN is able to achieve excellent transparency. Whereas our transparent cloak is specifically designed at 8 GHz, the simulation results show its working bandwidth ranges from about 7.5 to 9 GHz.

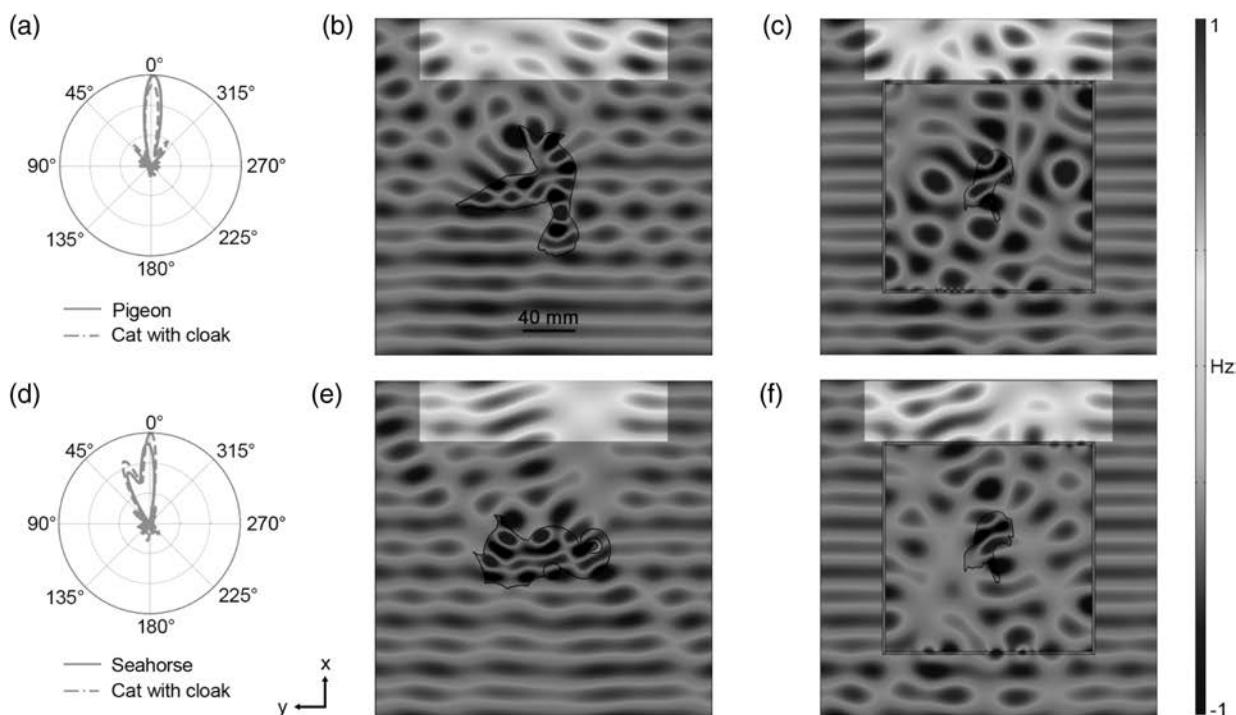
**Other functionalities designed by the T-NN.** In addition to the transparent invisibility cloak, the pretrained T-NN metasurfaces can also be used to meet many other user-oriented demands. As an example, we utilize the advances of the T-NN metasurfaces to mimic the EM characteristic of another object, for example, a pigeon or seahorse. To reach this goal, we firstly simulate the pigeon or seahorse model (with a relative



**Fig. 4.** Training results of the T-NN. (a) Learning curve of the T-NN for the far field, with an accuracy of 93.2%. To intuitively demonstrate the T-NN performance, we blindly select three RCS curves in the test set as the inputs [red curve in (c)], and output the metasurface arrangements from the intermediate layer, as shown in (b). In (c), we also plot the output of the T-NN (green curve), and the simulation result (purple curve) of the samples in (b). (d) Learning curve of the T-NN for the near field, with an accuracy of 92.4%; similar to the RCS above, we also blindly select three samples [upper part of (f)] as the inputs, and output the metasurface arrangements from the intermediate layer, as shown in (e). In (f), the output of the T-NN [middle part of (f)] and the simulation results [lower part of (f)] are also plotted. Obviously, these three in (f), as well as those in (c), are highly consistent with each other.



**Fig. 5.** Transparent invisibility cloak enabled by the pretrained T-NN. In an ideal case, the scattering of the transparent invisibility cloak should be zero, which is fed into the pretrained T-NN as the input. As such, we obtain the metasurface arrangements in (a). Based on (a), we obtain the normalized RCS of the dielectric cat with/without the cloak, as shown in (b). (c) and (d) are the simulated magnetic fields without/with the cloak, respectively, where the incident plane wave propagates from bottom to top. In (d), the field keeps almost flat after passing though the cloaking device, in stark contrast to that in (c).



**Fig. 6.** Other functionalities enabled by T-NN. (a) Normalized RCS of a dielectric pigeon and a dielectric cat surrounded by a bilayer metasurface; (b), (c) simulated magnetic field distribution of (b) the pigeon and (c) the cat with a bilayer metasurface; (d) normalized RCS of a dielectric seahorse and a dielectric cat with a bilayer metasurface; (e), (f) simulated magnetic field distribution of (e) the seahorse and (f) the cat with a bilayer metasurface.

permittivity of 3.5), and input the simulated far-field response [red lines in Figs. 6(a) and 6(d)] and near-field response [highlighted regions in Figs. 6(b) and 6(e)] into the pretrained T-NN. Then, according to the output of the T-NN, we resimulate the designed bilayer metasurface cloaks, as shown in Fig. 6. The high consistency between the ground truths and the T-NN predicted results suggests that our T-NN metasurfaces are generalized and applicable for different purposes.

### 3. CONCLUSION

In conclusion, we proposed an easy yet viable approach to realizing the transmitted metasurface cloak, and introduced a T-NN to efficiently streamline the inverse design process. The transmitted metasurfaces cloak hides a cat-shaped object sandwiched between two planar metasurfaces. The T-NN is deployed to address the hard-to-converge issue caused by the nonuniqueness phenomenon that widely exists in the EM inverse problem. Once pretrained, the T-NN can work for a customer-desired optical response in one single forward computation, including an invisibility cloak. Our work opens a new pathway to realizing a transparent cloak and enables a variety of other applications [20,33,34].

**Funding.** National Natural Science Foundation of China (11961141010, 61625502, 61975176, 61975182, 62071424); Top-Notch Young Talents Program of China; Fundamental Research Funds for the Central Universities.

**Acknowledgment.** C. Q., B. Z., and H. C. conceived the idea for this research. Z. Z. performed the simulation and built the neural network model. Z. Z. and C. Q. wrote the paper. Y. J. and Z. F. provided help with the deep neural network. E. L., R. H., T. C., B. Z., and H. C. shared their insights and contributed to discussions on the results. C. Q., R. H., H. C., and B. Z. supervised the project.

**Disclosures.** The authors declare no competing financial interests.

**Data availability.** The data that support the findings of this study are available from the authors upon reasonable request.

### REFERENCES

1. N. Engheta and R. W. Ziolkowski, *Metamaterials: Physics and Engineering Explorations* (Wiley, 2006).
2. J. B. Pendry, D. Schurig, and D. R. Smith, "Controlling electromagnetic fields," *Science* **312**, 1780–1782 (2006).
3. D. Schurig, J. J. Mock, B. J. Justice, S. A. Cummer, J. B. Pendry, A. F. Starr, and D. R. Smith, "Metamaterial electromagnetic cloak at microwave frequencies," *Science* **314**, 977–980 (2006).
4. Y. Lai, J. Ng, H. Y. Chen, D. Z. Han, J. J. Xiao, Z. Q. Zhang, and C. T. Chan, "Illusion optics: the optical transformation of an object into another object," *Phys. Rev. Lett.* **102**, 253902 (2009).
5. H. Y. Chen, C. T. Chan, and P. Sheng, "Transformation optics and metamaterials," *Nat. Mater.* **9**, 387–396 (2010).
6. Q. Chao, R. Li, Y. Jiang, B. Zheng, H. Wang, Z. Xu, and H. Chen, "Transient response of a signal through a dispersive invisibility cloak," *Opt. Lett.* **41**, 4911–4914 (2016).

7. B. Zheng, R. Zhu, L. Jing, Y. Yang, and H. Chen, "3D visible-light invisibility cloak," *Adv. Sci.* **5**, 1800056 (2018).
8. N. Yu, P. Genevet, M. A. Kats, and F. Aieta, "Light propagation with phase discontinuities: generalized laws of reflection and refraction," *Science* **334**, 333–337 (2011).
9. S. Sun, Q. He, S. Xiao, Q. Xu, X. Li, and L. Zhou, "Gradient-index meta-surfaces as a bridge linking propagating waves and surface waves," *Nat. Mater.* **11**, 426–431 (2012).
10. T. J. Cui, M. Q. Qi, X. Wan, Z. Jie, and C. Qiang, "Coding metamaterials, digital metamaterials and programmable metamaterials," *Light Sci. Appl.* **3**, e218 (2014).
11. T. Cai, G. M. Wang, S. W. Tang, H. X. Xu, J. W. Duan, H. J. Guo, F. X. Guan, S. L. Sun, Q. He, and L. Zhou, "High-efficiency and full-space manipulation of electromagnetic wave fronts with metasurfaces," *Phys. Rev. Appl.* **8**, 034033 (2017).
12. T. Cai, S. Tang, G. Wang, H. Xu, S. Sun, Q. He, and L. Zhou, "High-performance bifunctional metasurfaces in transmission and reflection geometries," *Adv. Opt. Mater.* **5**, 1600506 (2017).
13. C. Qian, X. Lin, Y. Yang, X. Xiong, H. Wang, E. Li, I. Kaminer, B. Zhang, and H. Chen, "Experimental observation of superscattering," *Phys. Rev. Lett.* **122**, 063901 (2019).
14. L. Li, H. Ruan, C. Liu, Y. Li, Y. Shuang, A. Alù, C. Qiu, and T. Cui, "Machine-learning reprogrammable metasurface imager," *Nat. Commun.* **10**, 1082 (2019).
15. C. Qian, X. Lin, X. Lin, J. Xu, Y. Sun, E. Li, B. Zhang, and H. Chen, "Performing optical logic operations by a diffractive neural network," *Light Sci. Appl.* **9**, 59 (2020).
16. X. Ni, Z. J. Wong, M. Mrejen, Y. Wang, and X. Zhang, "An ultrathin invisibility skin cloak for visible light," *Science* **349**, 1310–1314 (2015).
17. J. Zhang, Z. L. Mei, W. R. Zhang, Y. Fan, and T. J. Cui, "An ultrathin directional carpet cloak based on generalized Snell's law," *Appl. Phys. Lett.* **103**, 151115 (2013).
18. B. Orazbayev, N. M. Estakhri, M. Beruete, and A. Alu, "Terahertz carpet cloak based on a ring resonator metasurface," *Phys. Rev. B* **91**, 195444 (2015).
19. H. Lu, B. Zheng, T. Cai, C. Qian, Y. Yang, Z. Wang, and H. Chen, "Frequency-controlled focusing using achromatic metasurface," *Adv. Opt. Mater.* **9**, 2001311 (2020).
20. C. Qian, B. Zheng, Y. Shen, L. Jing, E. Li, L. Shen, and H. Chen, "Deep-learning-enabled self-adaptive microwave cloak without human intervention," *Nat. Photonics* **14**, 383–390 (2020).
21. H. Chu, Q. Li, B. Liu, J. Luo, S. Sun, Z. Hang, L. Zhou, and Y. Lai, "A hybrid invisibility cloak based on integration of transparent metasurfaces and zero-index materials," *Light Sci. Appl.* **7**, 50 (2018).
22. H. Li, M. Rosendo-López, Y. Zhu, X. Fan, D. Torrent, B. Liang, J. Cheng, and J. Christensen, "Ultrathin acoustic parity-time symmetric metasurface cloak," *Research* **2019**, 8345683 (2019).
23. Z. C. Liu, D. Y. Zhu, S. P. Rodrigues, K. Lee, and W. Cai, "Generative model for the inverse design of metasurfaces," *Nano Lett.* **18**, 6570–6576 (2018).
24. W. Ma, F. Cheng, and Y. M. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
25. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
26. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
27. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
28. L. Gao, X. Li, D. Liu, L. Wang, and Z. Yu, "A bidirectional deep neural network for accurate silicon color design," *Adv. Mater.* **31**, 1905467 (2019).
29. Y. Long, J. Ren, Y. Li, and H. Chen, "Inverse design of photonic topological state via machine learning," *Appl. Phys. Lett.* **114**, 181105 (2019).
30. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2020).
31. F. Liu, O. Tsilipakos, A. Pitilakis, A. C. Tasolamprou, M. S. Mirmoosa, N. V. Kantartzis, D. Kwon, J. Georgiou, K. Kossifos, M. A. Antoniades, M. Kafesaki, C. M. Soukoulis, and S. A. Tretyakov, "Intelligent metasurfaces with continuously tunable local surface impedance for multiple reconfigurable functions," *Phys. Rev. Appl.* **11**, 044024 (2019).
32. O. Tsilipakos, A. C. Tasolamprou, A. Pitilakis, F. Liu, X. Wang, M. S. Mirmoosa, D. C. Tzarouchis, S. Abadal, H. Taghvaei, C. Liaskos, A. Tsoliariadou, J. Georgiou, A. Cabellos-Aparicio, E. Alarcón, S. Ioannidis, A. Pitsillides, I. F. Akyildiz, N. V. Kantartzis, E. N. Economou, C. M. Soukoulis, M. Kafesaki, and S. Tretyakov, "Toward intelligent metasurfaces: the progress from globally tunable metasurfaces to software-defined metasurfaces with an embedded network of controllers," *Adv. Opt. Mater.* **8**, 2000783 (2020).
33. C. Qian, H. Wang, R. Li, B. Zheng, Z. Xu, and H. Chen, "Observing the transient buildup of a superscatterer in the time domain," *Opt. Express* **25**, 4967–4974 (2017).
34. S. Xu, F. Dong, W. Guo, D. Han, C. Qian, F. Gao, W. Su, H. Chen, and H. Sun, "Cross-wavelength invisibility integrated with various invisibility tactics," *Sci. Adv.* **6**, eabb3755 (2020).

# PHOTONICS Research

## Accurate inverse design of Fabry–Perot-cavity-based color filters far beyond sRGB via a bidirectional artificial neural network

PENG DAI,<sup>1,†</sup> YASI WANG,<sup>2,†</sup> YUEQIANG HU,<sup>2</sup> C. H. DE GROOT,<sup>1</sup> OTTO MUSKENS,<sup>3</sup>  HUIGAO DUAN,<sup>2,4</sup> AND RUOMENG HUANG<sup>1,5</sup> 

<sup>1</sup>School of Electronics and Computer Science, University of Southampton, Southampton SO17 1BJ, UK

<sup>2</sup>National Engineering Research Center for High Efficiency Grinding, College of Mechanical and Vehicle Engineering, Hunan University, Changsha 410082, China

<sup>3</sup>School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK

<sup>4</sup>e-mail: duanhg@hnu.edu.cn

<sup>5</sup>e-mail: r.huang@soton.ac.uk

Received 26 November 2020; revised 21 February 2021; accepted 1 March 2021; posted 1 March 2021 (Doc. ID 415141); published 30 April 2021

Structural color based on Fabry–Perot (F-P) cavity enables a wide color gamut with high resolution at submicroscopic scale by varying its geometrical parameters. The ability to design such parameters that can accurately display the desired color is therefore crucial to the manufacturing of F-P cavities for practical applications. This work reports the first inverse design of F-P cavity structure using deep learning through a bidirectional artificial neural network. It enables the production of a significantly wider coverage of color space that is over 215% of sRGB with extremely high accuracy, represented by an average  $\Delta E_{2000}$  value below 1.2. The superior performance of this structural color-based neural network is directly ascribed to the definition of loss function in the uniform CIE 1976-Lab color space. Over 100,000 times improvement in the design efficiency has been demonstrated by comparing the neural network to the metaheuristic optimization technique using an evolutionary algorithm when designing the famous painting of “Haystacks, end of Summer” by Claude Monet. Our results demonstrate that, with the correct selection of loss function, deep learning can be very powerful to achieve extremely accurate design of nanostructured color filters with very high efficiency. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415141>

### 1. INTRODUCTION

Structural color filters can display various colors by selectively transmitting or reflecting a specific wavelength by varying structural parameters rather than material components in the visible region [1]. They have received enormous interest recently due to their potential applications in chromatic display [2,3], color printing [4,5], optical encryption [6,7], solar cells [8], and so on. They largely exceed the conventional colorant-pigment-based filters in multiple aspects including nontoxicity, great scalability and durability, high resolution, and easy tenability [9]. Generation of structural color normally employs resonances (e.g., plasmonics, Mie scattering, guided mode resonance) from subwavelength patterns [10–13]. However, all these require nanoscale patterning, which involves complicated fabrication steps and can be cost prohibitive for high-volume and large-area applications.

Planar thin-film structures based on Fabry–Perot (F-P) cavity resonances are an alternative structural color filter technology based on the thin-film interference [14]. A typical F-P

resonator consists of a lossless dielectric that is sandwiched between two reflective metal layers. The function of color filtering is achieved by multiple round-trip phase delays of electromagnetic waves in the F-P resonator [15]. The resonant peak location can therefore be controlled by varying the dielectric layer thickness to obtain different colors. The full width at half-maximum (FWHM) of the peak, which relates to the color purity and brightness, can be tuned through the two metal layer thicknesses. Compared with the pattern-based filters, F-P-cavity-based structural color offers a much lower-costing and higher-scalability way of structural color manufacturing while offering larger color gamut and high color purity and contrast [16]. The high-index materials of phase compensation overlayers can also achieve great angle insensitivity for both transverse-magnetic (TM) and transverse-electric (TE) polarizations [17]. There are several works published on the aspect of improving the performance of F-P cavity-based structural color in both transmissive and reflective modes [18–20]. In addition, the F-P-cavity-based color filter also demonstrates high lateral

resolution. Wang *et al.* reported that its color crosstalk is comparable to those of plasmonic filters and pigmented filters with the same sizes [4]. A minimum pixel dimension of 500 nm (*ca.* 50,800 dpi) has also been successfully demonstrated, showing a great potential for F-P-cavity-based color filters to be applied in high-resolution colorization.

An important aspect for structural color is the ability to design a structure that can accurately display the desired color [21]. This conventional design task is realized by a trial-and-error method in which an initial random design is converged to the desired design through iterative optimization. Identification of a reasonable design therefore often requires an experienced designer with prior knowledge of the problem and a significant amount of calculations or simulations, which could be prohibitively slow as the complexity of the structure increases. However, this paradigm has recently been changed by the unprecedented development of deep learning techniques [21–23].

Deep learning is an important subset of machine learning in which multilayered artificial neural networks are utilized to achieve high-accuracy predictions and classifications [24]. Several works have been reported recently in which neural networks were used to inversely design the nanostructure and material to achieve desired optical responses [25–31]. Before the networks can perform the intended functions, a training process needs to take place in which a dataset of structural parameters to color relations is required. This dataset normally involves a large amount of relations, and it needs to be generated by theoretical calculation or simulation. However, this is a one-time investment, and no more computational resources will be consumed once the network is properly trained. Pioneering works have also been reported on deep-learning-aided structural color design [32–36]. Hemmatyar *et al.* designed and optimized the hafnia array based all-dielectric metasurfaces via neural network to generate a wide color gamut [32]. Gao *et al.* reported a structural color inverse design by employing a bidirectional neural network [35], which was first proposed by Liu *et al.* [37]. The model consists of a fully connected inverse network that directly connects to a pretrained forward network. The inverse network can produce the design of the structural parameters for desired colors, while the forward network predicts the colors from the structural parameter inputs. The training of the inverse network is conducted by feeding the inversely designed structures directly into the pretrained forward network, and the network is optimized by minimizing the difference between the predicted color and the input color through backpropagation. This model has been widely used in different nanophotonic applications, as it overcomes the issue of nonuniqueness [31,38,39].

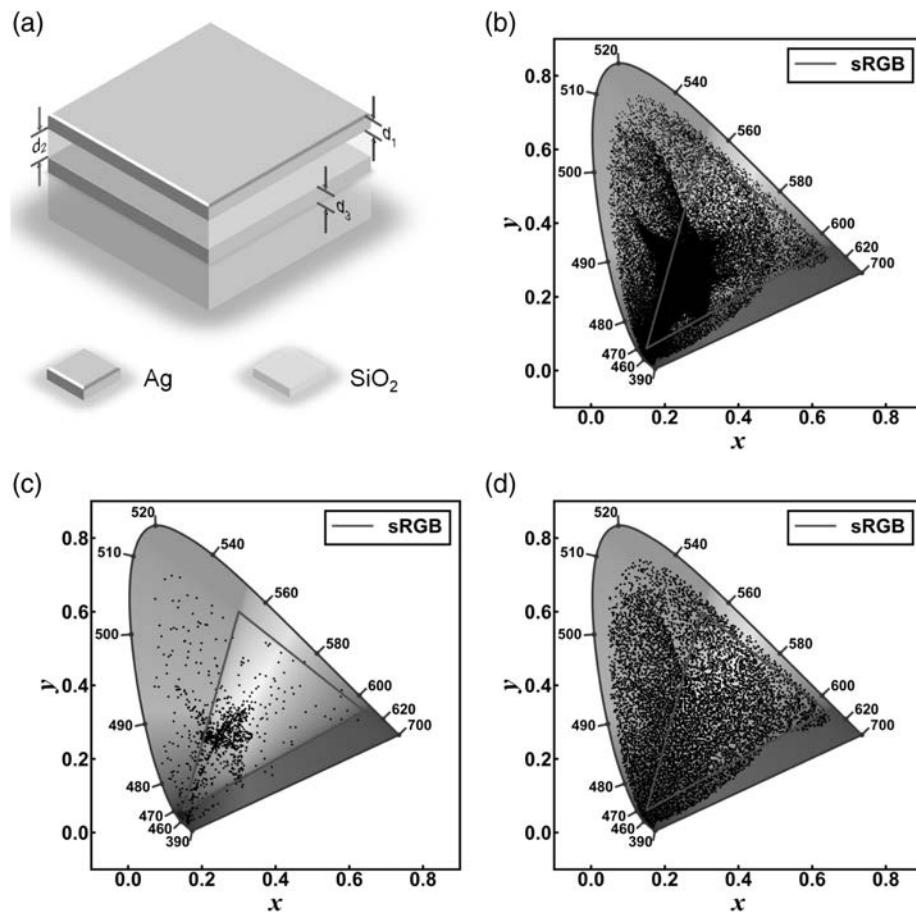
Although previous works have achieved remarkable progress in the structural color inverse design, one remaining challenge is the relatively low design accuracy (large color differences between design and target) even when extremely low validation loss is achieved. One key reason for that is the nonuniformity of the color space from which the loss function was defined. The loss function measures the difference between the network predicted value and the true value in the dataset. It is crucial for neural networks, as the training is a process of minimizing the

loss function [40]. An unsuitable selection of loss function may lead to a very low loss function but relatively higher prediction error [41]. Defining loss function in a nonuniform color space such as CIE 1931-XYZ means the same Euclidean distance among the **XYZ** vectors may signify different color differences, resulting in biased optimization against some colors [42].

Here we report, to the best of our knowledge, the first inverse design of F-P cavity structure using deep learning through a bidirectional artificial neural network. This structure enables the production of a significantly wider color gamut that is over 215% of sRGB color space, which is essential for real display applications. By defining the training loss function in the uniform CIE 1976-Lab color space where the same Euclidean distance indicates same color difference, our network is able to achieve a much higher design accuracy with average color difference  $\Delta E_{2000}$  below 1.2. The high design efficiency of the network is also evaluated by comparison to the evolutionary algorithm, which shows over 100,000 times savings of computational resources. The demonstration of this work offers exciting prospects for deep learning techniques to be used to achieve accurate and efficient structural color designs for a wide range of different applications.

## 2. DATASET GENERATION

The schematic diagram of the transmissive F-P-cavity-based color filter employed in this work is illustrated in Fig. 1(a). A transmissive-type of color filter was chosen because of its wide application in spectrometers, CMOS image sensors, and liquid crystal displays [43–45]. Compared with reflective filters, transmissive filters enjoy the advantages of single-mode operation, which could lead to a higher color purity. However, it is worth mentioning that a similar approach can also be used on the reflective type. The system has a trilayer metal-insulator-metal (MIM) films stack on the quartz substrate in which a  $\text{SiO}_2$  dielectric layer is sandwiched between two Ag metal layers. Optical interference occurs when white light enters the cavity which filters out the wavelengths that are not matching the resonant wavelength of this multilayer system. With the materials of the dielectric and metal layers fixed, the most critical parameters here affecting the resonant wavelengths are the thickness of each layer, which are represented as  $d_1$ ,  $d_2$ , and  $d_3$  as shown in Fig. 1(a). In this work, the ranges of  $d_1$ ,  $d_2$ , and  $d_3$  are set to be 0 to 50 nm, 0 to 1000 nm, and 0 to 50 nm to allow a large color gamut coverage. Only integer values are selected in this work to ensure the compatibility with fabrication techniques. A total of 101,000 parameter combinations are randomly generated, and the corresponding transmissive spectra from 380 to 780 nm are computed by the multiple beam interference formulas [16]. However, color is not a property of electromagnetic radiation but a subjective perception of an observer. Color-matching functions are then required to convert the transmissive spectra into corresponding color coordinates in CIE 1931-XYZ color space to generate a dataset containing 101,000 parameters **D** ( $d_1, d_2, d_3$ ) to color **XYZ** ( $X, Y, Z$ ) relations. This dataset is divided into three groups for training (90,000), validation (10,000), and testing (1000) purposes. All colors generated in the training and validation dataset are plotted in the CIE 1931-*xy* chromaticity diagram in



**Fig. 1.** (a) Schematic illustration of a transmissive F-P-cavity-based color filter with an MIM structure. (b) The training and validation dataset generated by the F-P cavity plotted in a CIE 1931-xy chromaticity diagram. The red triangle depicts the boundary of sRGB color space for a guide of the eye. (c) The original 1000 testing data and (d) the 7000 uniformly distributed testing data plotted in the CIE 1931-xy chromaticity diagram.

Fig. 1(b). It is clear that our Ag-SiO<sub>2</sub>-Ag F-P-cavity-based color filter in this work can achieve a substantially larger (*ca.* 215%) color gamut than the sRGB color space (plotted by red lines for reference). The large gamut of coverage is one of the advantages of F-P-cavity-based structural color and is particularly beneficial for real applications such as display and full-color nanoprinting [4,19]. By varying the thickness of the dielectric layer ( $d_2$ ), the transmission peak could be swept across the whole color in the visible light range. The modification of the Ag layer thicknesses ( $d_1$  and  $d_3$ ) serve to further tune the FWHM of the transmittance peak, resulting in a large gamut of coverage. The testing dataset is plotted in Fig. 1(c) (referred to as the testing set below). To further test the robustness of our networks, we also generated an additional testing set with 7000 colors that are uniformly distributed on the CIE 1931-xy chromaticity diagram [shown in Fig. 1(d) and referred to as the uniform testing set below].

### 3. FORWARD NEURAL NETWORK CONSTRUCTION

A forward neural network (FNN) was first trained to obtain accurate prediction of colors based on the layer thicknesses. Prior to training, a loss function needs to be established.

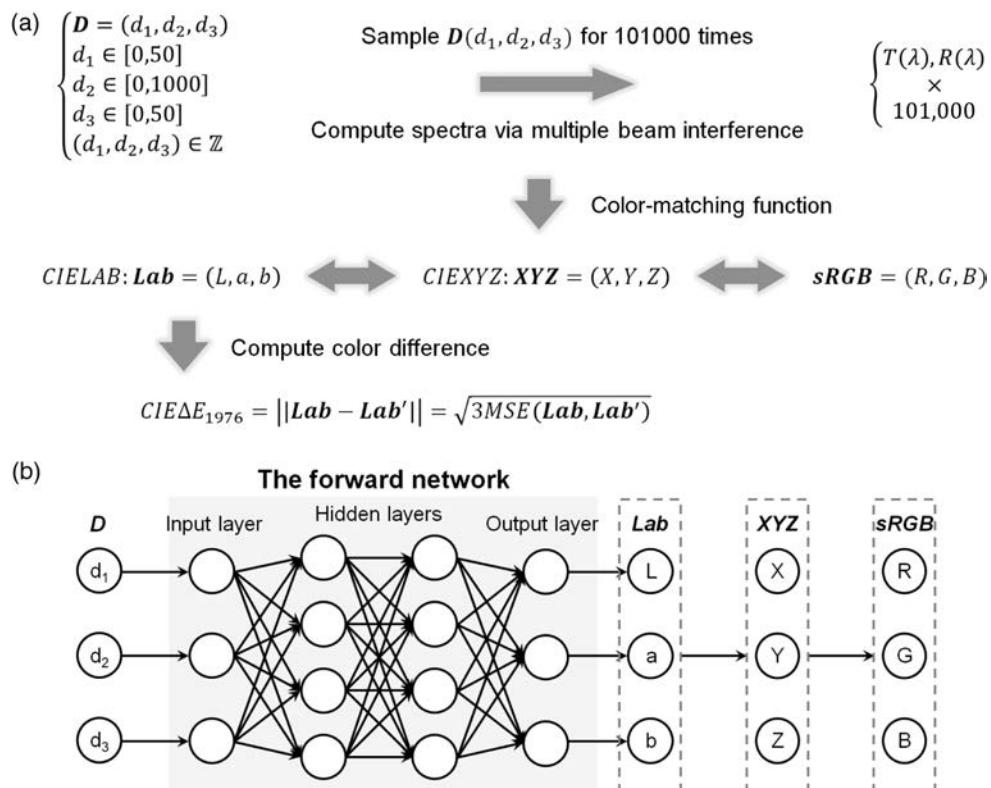
Although our original dataset converts the spectrum into a CIE 1931-XYZ tristimulus vector, it is not a suitable output for loss function definition due to its nonuniformity [42]. sRGB color space is also not ideal, as the conversion between XYZ and sRGB is not reversible when the color is outside of the sRGB color space. On the other hand, the CIE 1976-Lab color space has a one-to-one correspondence to the CIE 1931-XYZ but with much better uniformity, rendering it a more suitable color space for accurate color difference identification. In fact, the color difference function CIE  $\Delta E_{1976}$  is defined by the Euclidean distance of two Lab vectors ( $L, a, b$ ). This property is particularly beneficial in neural network training. By defining the loss function to be mean squared error (MSE) between the predicted and original Lab values, it can be directly converted to the actual color difference ( $\Delta E_{1976}$ ) and enable higher accuracy. Figure 2(a) shows the summary of the dataset preparation process for the FNN. After obtaining XYZ from the spectrum, it was converted to Lab, which was then used to construct the loss function and identify color difference. The architecture of the FNN is illustrated in Fig. 2(b) and is composed of a fully connected neural network (NN) including one input layer, one output layer, and several hidden layers. It takes the parameter  $\mathbf{D}$  as input and outputs the Lab, which can be converted to other color vectors such as XYZ and sRGB for different applications.

During the deep learning process, the training data group is fed to the FNN to continuously adjust the weight and bias of each connection with every batch of data in epochs. This is achieved by backpropagation of the loss function. The selection of the hyperparameters (i.e., number of hidden layers and neurons per layer) is crucial to the performance of the network [21]. A systematic study was therefore conducted to investigate the impact of hyperparameters for this FNN. The CIE  $\Delta E_{2000}$  color difference was chosen here to provide a better quantification of the FNN performance. Similar to the  $\Delta E_{1976}$ , the  $\Delta E_{2000}$  color difference is also a function of the two **Lab** values but corresponds better with the way in which human observers perceive small color differences and hence is used as the metric for design accuracy in this work [46]. The  $\Delta E_{2000}$  can be classified into five groups: 1)  $\Delta E_{2000} < 1$ , it can be considered no color difference; 2)  $1 < \Delta E_{2000} < 2$ , the difference can be observed by experienced persons; 3)  $2 < \Delta E_{2000} < 3.5$ , the difference can be observed by unexperienced persons; 4)  $3.5 < \Delta E_{2000} < 5$ , a clear difference can be noticed; and 5)  $5 < \Delta E_{2000}$ , two different colors are observed [47]. The distribution of  $\Delta E_{2000}$  from the testing set of each FNN is plotted in Fig. 3(a) as a function of layer number. It provides a clear indication on the color prediction performance of each FNN. The average  $\Delta E_{2000}$  values (blue squares) can therefore be used to provide a quantified metric for the performance. It can be observed that the average  $\Delta E_{2000}$  in the testing set decreases from 1.87 with two hidden layers to 0.44 with seven hidden

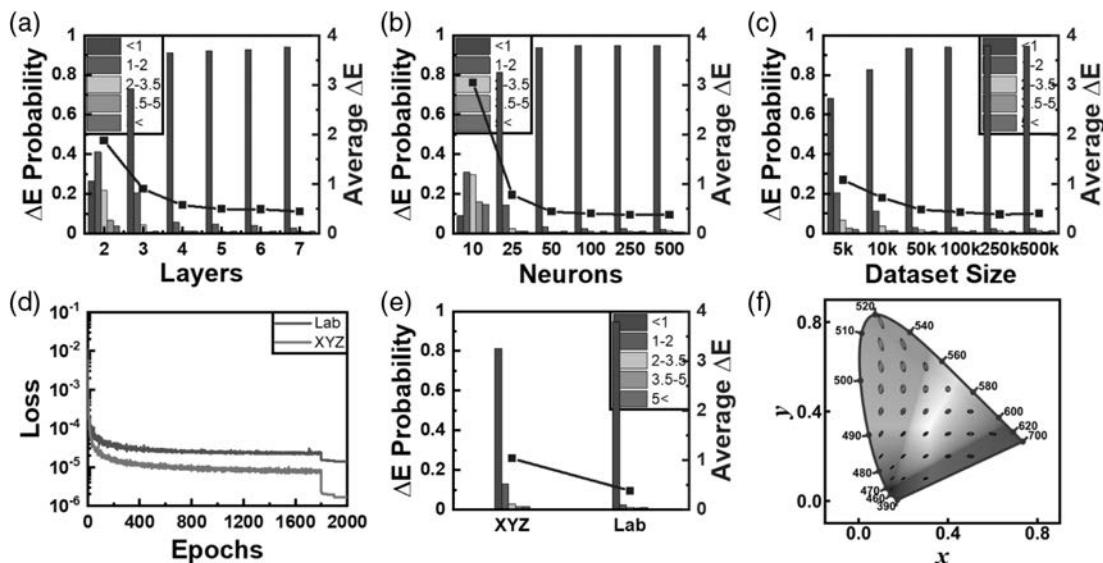
layers, indicating that the FNN with seven hidden layers has the best performance in our optimization range. The impact of number of neurons per layer was also investigated while the number of layers was fixed at seven.

Similar to the optimization of hidden layer numbers, the  $\Delta E_{2000}$  distribution and average values are plotted in Fig. 3(b). The average  $\Delta E_{2000}$  in the testing set plunged sharply from 3.05 to 0.44 with the increase of neuron number from 10 to 50 and subsequently bottomed out at 0.38 for 250 neurons. Therefore, the architecture of the FNN is optimized with seven hidden layers and 250 neurons in each hidden layer. The influence of the dataset size was also investigated as shown in Fig. 3(c). The results suggest that sufficient dataset size (over 50,000 in our case) is required for the network to achieve a good performance of an average  $\Delta E$  value below 0.5 in our case.

The high accuracy of our FNN is further confirmed by evaluation using the uniform testing set, resulting in a  $\Delta E_{2000}$  of 0.35. We believe this high accuracy can be ascribed to the use of uniform CIE 1976-Lab color space for the network training. An FNN with the same network architecture using **XYZ** as the output was also trained for comparison as shown in Figs. 3(d) and 3(e). Although a lower MSE (*ca.*  $10^{-6}$ ) was achieved by the FNN with **XYZ** output, the  $\Delta E_{2000}$  distribution is significantly poorer than the one with **Lab** output with a decrease of  $\Delta E_{2000}$  values observed from 1.03 to 0.38. To better explain the advantage of our approach, Fig. 3(f)



**Fig. 2.** Forward neural network for predicting F-P cavity structural colors. (a) The relationships of different parameters in the dataset. The single arrow means the transformation is unidirectional, the double arrow means the transformation is reversible. (b) The architecture of the forward neural network with input layer of geometric parameter **D**, hidden layers, and output layers of Lab color values.



**Fig. 3.** Forward neural network training for predicting F-P cavity structural colors. The histogram of the probability and average values of  $\Delta E_{2000}$  of the FNNs with different (a) hidden layer number, (b) neuron number per layer, and (c) the FNN with seven hidden layers and 250 neurons in each hidden layer for different dataset size. (d) The training loss curves for defining the loss function in CIE 1931-XYZ and CIE 1976-Lab color spaces. (e) The probability histogram and average values of  $\Delta E_{2000}$  comparisons while the loss function is defined in CIE 1931-XYZ color space and CIE 1976-Lab color space. (f) The selected colors in the CIE 1931-xy chromaticity diagram with the boundary of each ellipse representing the colors that have a  $\Delta E_{1976}$  of 6 to the selected color.

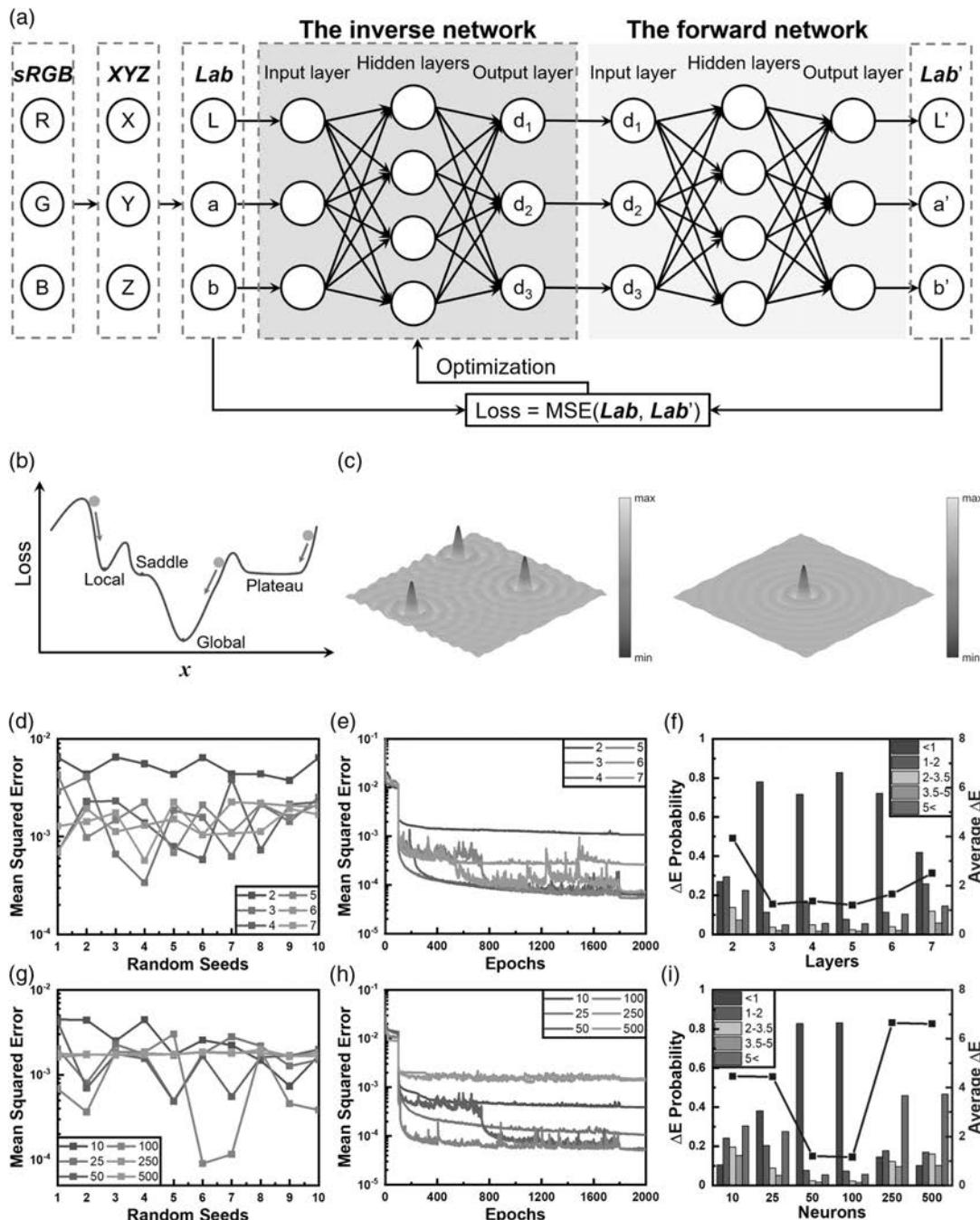
presents a selection of colors in the CIE 1931-*xy* chromaticity diagram with the boundary of each ellipse representing the colors that have a  $\Delta E_{1976}$  of 6 to the selected color. It is clear that the ellipses in the green and red regions are larger than that in the blue region due to the nonuniformity of the CIE 1931-XYZ color space. One could obtain a very small  $\Delta XYZ$  in the green region, but the actual reduction of color difference might be limited. This suggests that minimizing  $\Delta XYZ$  values, especially in the green and red regions, will not be as effective in reducing color difference as the approach suggested in our work. This proves that the correct selection of loss function is critical to the performance of the neural network.

#### 4. INVERSE NEURAL NETWORK CONSTRUCTION

The training of the inverse neural network (INN) is more challenging due to the nonuniqueness nature that one color can be formed by different F-P cavity structures. This multisolution property could lead to the adjustment of weight being pulled to different local or global minima during the training process, making the training difficult to converge. The bidirectional neural network architecture and tandem training strategy were therefore employed in this work in which the output parameter  $\mathbf{D}'(d'_1, d'_2, d'_3)$  from the INN was directly fed into our pre-trained FNN to generate predicted  $\mathbf{Lab}'(L', a', b')$  as shown in Fig. 4(a). The parameters of FNN are fixed during the INN training. The loss function can be defined as the MSE between the original and predicted Lab vectors instead of parameter  $\mathbf{D}$ . This means the INN will be optimized to match the desired color rather than the structure, avoiding the non-uniqueness problem in the training process. Here we adopted a

recently reported penalized tandem training strategy to further enhance the robustness of the network in the initial 100 epochs [39], which indicates that the inverse error (the MSE between the original  $\mathbf{D}$  and predicted  $\mathbf{D}'$ ) was included in the loss function for the first 100 epochs to avoid the predictions violating the ground truth of  $\mathbf{D}$ .

Another important factor in INN training is the selection of a random seed. The initial weights are obtained from random sampling using determined distributions (e.g., Xavier or Kaiming initialization). The training is therefore repeatable once the selection of a random seed was fixed. Different random seeds could place the INN at different starting positions in the loss plane [Fig. 4(b)]. In the case of multiple solutions, the loss plane is very complicated due to the interference of different global or local minima [Fig. 4(c), left]. Different starting positions could cause the INN to converge into a different local/global minimum, or in the case of a poor initialization, struggle to even converge. Selection of the random seed therefore plays a vital role in the INN training process. In this work, each INN underwent a process of random seed selection before hyperparameter optimization took place. The summary of the MSEs after 200 training epochs for INNs with varying hidden layers is presented in Fig. 4(d). It is obvious that the training progress differs significantly due to the selection of different random seeds, and different INN architectures prefer different random seed groups for optimized performance. This is particularly important for multisolution questions with the existence of a large number of local and global minima. It is less critical for single-solution questions (e.g., the FNN in this work), as the loss plane is relatively simple for the network to converge [Fig. 4(c), right] regardless of the initial starting position.



**Fig. 4.** Inverse neural network for predicting F-P cavity structural colors. (a) The bidirectional architecture with input layer of Lab values and output layer of geometric parameter  $\mathbf{D}$  and connected to the pretrained forward neural network. (b) The schematic of different weights' initialization positions. (c) The loss surface schematics of nonuniqueness (left) and uniqueness (right) problems, respectively. The MSEs after 200 epochs as a function of random seed, the training loss curves, the histogram of the distribution, and average values of  $\Delta E_{2000}$  of the INNs with (d)–(f) different numbers of hidden layers and (g)–(i) different numbers of neurons per layer.

Figures 4(e) and 4(f) present the optimization of hidden layer numbers for INN. A drop of MSE at the 100th epoch in each training is due to the removal of the inverse error term in the loss function as mentioned above. Unlike the FNN where higher network complexity always leads to a better performance in the training set, a reduction of performance was observed for the INN with more than five layers. This inferior

performance for more complicated INNs may be caused by the increased dimensions in the loss plane and number of existing global minima, which causes the network convergence to be more difficult. Similar behavior was also observed in the process of neuron number optimization as shown in Figs. 4(g)–4(i). A final INN with five hidden layers with 100 neurons was identified to be the best INN architecture to be used in this work.

This network enables the average  $\Delta E_{2000}$  of 1.16 in the non-uniform testing set and 1.18 in the uniform testing set, which are superior to most of the commercial display equipment. For example, the high-end Dell UltraSharp 32 PremierColor UltraHD 8K Monitor is factory calibrated at 100% sRGB coverage to an accuracy of  $\Delta E_{2000}$  less than 2.

## 5. PERFORMANCE EVALUATION

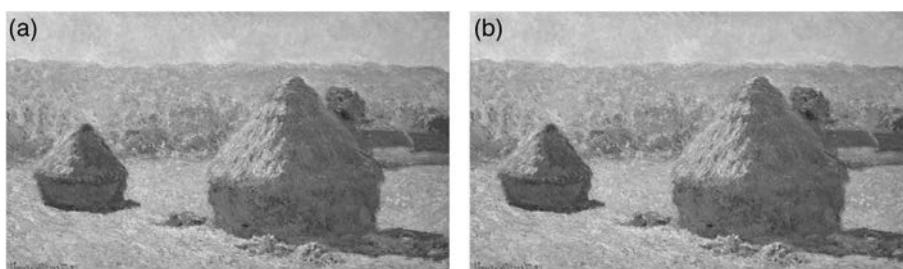
To better evaluate the performance of our INN, we randomly selected 6 F-P cavity structures (**D** values) that the INN has never seen before, as listed in Table 1. These structures were converted to the corresponding spectra and subsequently **Lab** values as discussed previously. The target **Lab** values were then fed into our INN to obtain the designed structures **D'**. Further tests by converting the designed structures to designed colors through the theoretical calculation have resulted in very close matches with the targeted colors. The color differences, represented by the  $\Delta E_{2000}$  values, are below 1 in all six cases, suggesting that human eyes are not able to distinguish their differences. More importantly, in five of the six cases, our

INN has produced designed structures that are significantly different from the original to the targeted structures while obtaining small color difference. This demonstrates the existence of nonuniqueness solutions for INN. More importantly, it highlights the ability of a neural network to discover solutions outside of the boundaries of the training data. Nanophotonic research has become more computation intensive due to the large spatial degrees of freedom and wide choice of materials [22]. Such ability to identify new design structures that can never be found through a conventional forward design technique would be extremely beneficial for discovering novel findings in nanophotonics.

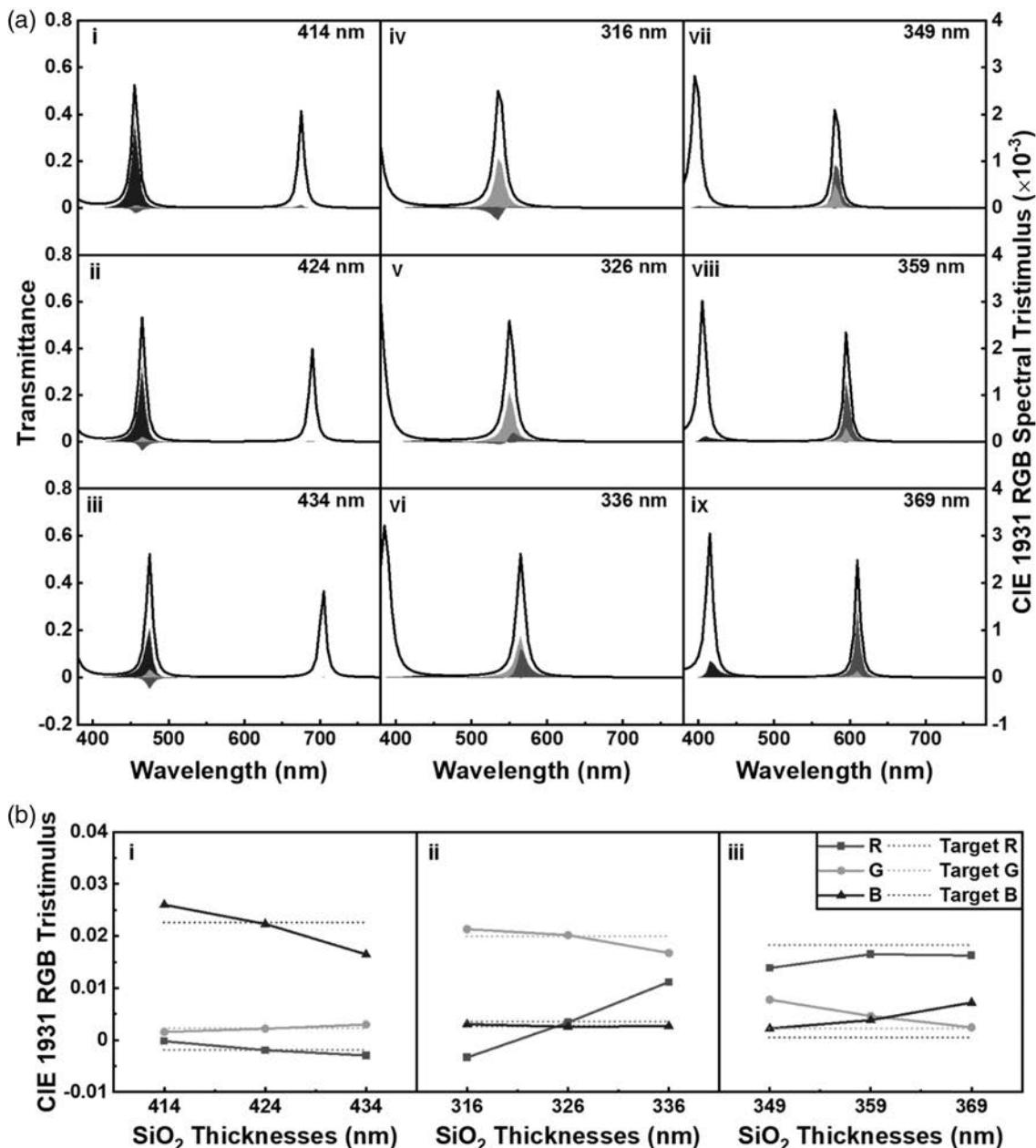
The excellent performance of our network is further supported by designing the F-P cavity structures to reproduce the painting of "Haystacks, end of Summer" by Claude Monet. This was done by extracting the **sRGB** values of all  $2000 \times 1176$  (2,352,000) color pixels from the original painting [shown in Fig. 5(a)] and inputting them into the bidirectional network. The INN outputs the designed geometric parameters for all 2,352,000 pixels, and the reconstructed paintings were subsequently generated and plotted in Fig. 5(b).

**Table 1. Comparison of the Six Randomly Selected Target Structure D, Lab, and sRGB Color Values with the Corresponding Values Inversely Designed by the Inverse Neural Network**

Target D (nm)	Target sRGB		Target Lab	Design D (nm)	Design Lab	Design sRGB	$\Delta E$
31-365-31	142-27-80 142-27-80		L: 32.0 a: 50.2 b: -1.7	32-365-30	L: 31.8 a: 49.6 b: -1.4	141-28-79 141-28-79	0.27
44-152-11	95-67-59 95-67-59		L: 31.2 a: 10.7 b: 11.7	43-345-15	L: 31.5 a: 10.4 b: 12.4	96-68-55 96-68-55	0.70
29-141-12	163-142-97 163-142-97		L: 59.8 a: 1.7 b: 26.8	25-336-15	L: 60.0 a: 1.6 b: 27.5	164-142-96 164-142-96	0.37
29-121-26	(-600)-165-114 0-165-114		L: 56.4 a: -84.2 b: 11.0	23-293-20	L: 56.5 a: -84.8 b: 11.5	(-605)-165-113 0-165-113	0.25
20-385-44	89(-44)-110 89-0-110		L: 16.9 a: 61.8 b: -46.0	41-387-23	L: 16.9 a: 62.0 b: -45.7	89(-45)-109 89-0-109	0.18
19-839-18	44-140-166 44-140-166		L: 54.2 a: -19.9 b: -21.6	24-271-5	L: 54.3 a: -19.1 b: -21.2	44-141-166 44-141-166	0.29



**Fig. 5.** Actual application of the INN-based structural color design. (a) The origin and (b) the reproduction via INN of the painting "Haystacks, end of Summer" by Claude Monet. Reproduction of "Haystacks, end of Summer" is permitted by the Musée d'Orsay, Paris (RF 1975 3).



**Fig. 6.** Transmissive spectra and corresponding CIE 1931-RGB tristimulus values for the designed colors. The transmissive spectra (black line) and the contribution from the three stimuli (shades underneath the line) for the (i, ii, iii) blue, (iv, v, vi) green, and (vii, viii, ix) red color designed by the INN in this work. Within each figure, the middle row figure presents the original design, while the top and bottom figures represent the spectra from a 10 nm thinner layer and a 10 nm thicker dielectric layer. (b) The CIE 1931-RGB tristimulus values as a function of dielectric layer thickness for the blue, green, and red colors, respectively. The CIE 1931-RGB tristimulus values of the targeted colors are also included (dotted lines) for comparison.

An extremely high accuracy can be clearly observed, as the difference between the two images is almost undistinguishable. This proves the robustness and accuracy of our network in designing colors over a wide gamut.

We will now provide a detailed analysis of the network performance by evaluating the spectra of the blue (sRGB 0, 0, 102), green (sRGB 0, 102, 0), and red (sRGB 102, 0, 0) color filters designed by our INN. The designed F-P cavity for blue

color has a dielectric thickness ( $d_2$ ) of 424 nm, and its transmissive spectrum is characterized by a main peak at 465 nm with a secondary peak at 690 nm as shown in Fig. 6(a-ii). The contribution from each CIE 1931-RGB spectral tristimulus is demonstrated by the size of the shades underneath the spectrum. It is obvious that the designed blue color consists of a majority of blue stimuli with a small proportion of green and red stimuli, the integrals of which are extremely close to the

definition of blue color in the sRGB color space [shown by the dashed line in Fig. 6(b-i)], and the corresponding CIE 1931-RGB tristimulus values can be obtained by converting the CIE 1931-XYZ tristimulus]. Almost identical values are obtained with a  $\Delta E_{2000}$  of 0.07. An unoptimized design with a dielectric layer 10 nm thinner will cause a blueshift of the spectrum [Fig. 6(a-i)]. This results in an increased contribution from the red stimuli but a reduced contribution from the green stimuli, driving the designed color away from the blue definition with a  $\Delta E_{2000}$  of 6.16. Similarly, a 10 nm thicker dielectric layer induces a redshift with more contribution from green stimuli and less contribution from red stimuli, resulting in a  $\Delta E_{2000}$  of 11.83. Similar behavior is also observed on the design of green color as shown in Fig. 6(a-iii). Our design with a dielectric thickness ( $d_2$ ) of 326 nm results in a  $\Delta E_{2000}$  of 0.35 [Figs. 6(a-v) and 6(b-ii)], whereas a  $\pm 10$  nm change of the thickness alters the contribution from the CIE 1931-RGB tristimulus and results in inferior color design with  $\Delta E_{2000}$  of 7.72 ( $d_2 - 10$  nm) or 11.83 ( $d_2 + 10$  nm) as shown in Fig. 6(a-iv/vi). It is clear that our network is able to identify the optimized thickness that places the spectrum in the right position to achieve minimum difference from the desired color.

It is also worth pointing out the limitation of our network, which is manifested in the design of the red color. The network selects a design with  $d_2$  of 369 nm, which uses the second-order peak at 595 nm, resulting in a  $\Delta E_{2000}$  of 8.93 [Fig. 6(a-viii)]. Although it outperforms the other two scenarios with  $\pm 10$  nm of  $d_2$  [ $\Delta E_{2000}$  of 20.21 and 19.64, respectively, shown in Fig. 6(a-vii/ix)], it fails to select the first-order peak, which could realize a much better design by using a smaller thickness ( $d_2 \sim 160$  nm). This suggests that the network is not flexible enough when a large change of thickness is required. This may be attributed to two factors. The first factor is the uneven color distribution of the training dataset caused by the nonlinear relation among thickness, spectrum, and color. The resonant cavity lengths in the F-P cavity are close to the positive integer times of  $\lambda/4n$ , where  $n$  is the refractive index of the dielectric layer and  $\lambda$  is the light wavelength in free space. The wavelengths of blue ( $\sim 490$  nm) and green ( $\sim 550$  nm) are shorter than that of the red ( $\sim 610$  nm). Higher-order blue and green peaks therefore appear more frequently than red within the same dielectric layer range. Hence, colors near the red region are significantly under-represented. This leads to the network being trained in favor towards optimization of colors in the blue and green regions. The second factor lies in the inherent limitation with tandem network architecture, which suffers from mode collapse [23]. The high-quality first-order red peak was abandoned by the network during the training process to achieve a high overall accuracy for all colors. This results in the red colors being predicted to the higher orders, leading to limited quality. These two factors are believed to contribute to the difficulty of designing colors in the red region, and further improvements in both the network design and training process are required to tackle this challenge.

Finally, we evaluate the computational efficiency of our network by comparing it with an evolutionary algorithm—a popular metaheuristic optimization technique [48]. Compared with a neural network, an evolutionary algorithm performs better in

the design of red color (details of EA can be found in Section 7). This forward design process was able to find the first-order red peak, achieving a  $\Delta E_{2000}$  of 0.73. The designs of green and blue colors are not as good as our INN with  $\Delta E_{2000}$  of 1.36 and 1.11, respectively. Moreover, this method demands much more computational resources than INN and is unpractical in real applications. For example, the time required to design the painting of “Haystacks, end of Summer” (500 × 297, 148,500 pixels) via our INN was 0.17 s. It took the evolutionary algorithm 4.8 h to design the same number of pixels under the same computational environment (see Section 7, Method). This translates to over 100,000 times of savings in time and computational resources for our network. In addition, the overall design accuracy for those 148,500 colors obtained from the network is significantly better ( $\Delta E_{2000}$  of 0.78) than that designed by EA ( $\Delta E_{2000}$  of 1.18), further proving the superior performance in both computational efficiency and design accuracy of our network.

## 6. CONCLUSION

In conclusion, we demonstrate the use of a bidirectional neural network to inversely design the geometric structures of F-P cavity color filters. This work leads to a gamut coverage that is 215% of the sRGB color space. By selecting the uniform CIE 1976-Lab color space over the conventional CIE 1931-XYZ color space as the representation of color, the bidirectional network has shown a superior accuracy for color design with an average  $\Delta E_{2000}$  value below 1.2 in the testing set. This excellent performance is also verified by comparison with the gradient-free evolutionary algorithm in which our network demonstrates a 100,000 times design efficiency improvement with higher accuracy when designing the famous painting “Haystacks, end of Summer” by Claude Monet. The challenges in designing colors at longer wavelength due to uneven dataset distribution and the continuous gradient descent nature of artificial neural networks are also discussed. This proposed model will contribute to the establishment of standard procedure for future design of nanostructured color filters with deep learning technology.

## 7. METHOD

**Data Processing.** Before the neural networks are trained, the datasets are normalized from 0 to 1. By doing this, the effects of unit, magnitude, and dimension of the dataset are waived, which is able to improve the performance and probability of convergence. This process was done by the open-source machine learning library *Scikit-Learn*.

**Color Conversion.** The color conversions, which include the conversions between different color spaces (e.g., sRGB to CIE 1931-XYZ and CIE 1931-XYZ to CIE 1976-Lab) and color difference ( $\Delta E_{2000}$ ), were performed by the open-source library *Colour-Science*.

**Deep Learning.** All the deep learning models and training were developed and performed on the open-source deep learning framework *PyTorch*.

**Training Hyperparameters.** The training hyperparameters are listed as follows. Epochs: 2000; batch size: 64; activation function: ReLU; loss function: mean squared error (MSE); optimizer: Adam; learning rate: 0.001; learning rate scheduler:

MultiStepLR; milestones = [1800,1900] (forward training) and [1,1800,1900] (inverse training); gamma = 0.1. The Kaiming uniform initialization method was adopted in this work to investigate the impact of random seed selection [49].

**Loss Functions.** In the training of the forward neural network, the loss function is defined as  $\text{Loss}_{\text{forward}} = \text{MSE}(\text{Lab}_{\text{predicted}}, \text{Lab}_{\text{truth}})$ . In the training of the inverse neural network, a penalty term in the loss function has been introduced for the first 100 epochs to ensure  $D_{\text{predicted}}$  is not too far away from the input dimension  $D_{\text{truth}}$  (i.e., reduce the chance of generating negative values). The penalized loss function is defined as  $\text{Loss}_{\text{inverse}} = \text{MSE}(\text{Lab}_{\text{predicted}}, \text{Lab}_{\text{truth}}) + 0.2 \times \text{MSE}(D_{\text{predicted}}, D_{\text{truth}})$ . The second term in the equation is a penalty term, which is used to control the inverse output error compared with the ground truth of the geometric parameters. The penalty coefficient  $\lambda$  is set to be 0.2 in our case. After the 100 epochs, the loss function is defined as  $\text{Loss}_{\text{inverse}} = \text{MSE}(\text{Lab}_{\text{predicted}}, \text{Lab}_{\text{truth}})$ .

**Evolutionary Algorithm.** The evolutionary algorithm (EA) is one type of gradient-free optimization method, and it is an appealing option for solving this optimization problem. It uses a refined iterative process in which an elite percentage of the individuals are retained through each iteration, allowing the samples to genetically evolve until the best option has been identified. The evolutionary algorithm in this work was realized by open-source EA library **DEAP**. The EA generation and population applied in this work are 100 and 50, respectively. The EA designed geometrical parameters ( $d_1, d_2, d_3$ ) are (39 nm, 163 nm, 46 nm) for red (102, 0, 0); (33 nm, 138 nm, 50 nm) for green (0, 102, 0); and (21 nm, 417 nm, 49 nm) for blue (0, 0, 102). The  $\Delta E_{2000}$  are 0.73, 1.36 and 1.11, respectively.

**Computational Environment.** CPU: Intel Core i9-9900K; GPU: nVIDIA RTX 2070; RAM: 48 GB; OS: Windows 10 Pro; Python version: 3.7. The same computational environment was used for the ANN and EA methods in this work.

**Funding.** International Exchange Scheme (IEC\NSFC \170193) between Royal Society (UK) and the National Natural Science Foundation of China (China).

**Acknowledgment.** All data supporting this study are openly available from the University of Southampton repository at DOI: <https://doi.org/10.5258/SOTON/D1686>.

**Disclosures.** The authors declare no conflicts of interest.

<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

- N. Dean, "Colouring at the nanoscale," *Nat. Nanotechnol.* **10**, 15–16 (2015).
- J. Hong, E. Chan, T. Chang, T.-C. Fung, B. Hong, C. Kim, J. Ma, Y. Pan, R. Van Lier, S. Wang, B. Wen, and L. Zhou, "Continuous color reflective displays using interferometric absorption," *Optica* **2**, 589–597 (2015).
- L. Shao, X. Zhuo, and J. Wang, "Advanced plasmonic materials for dynamic color display," *Adv. Mater.* **30**, 1704338 (2018).
- Y. Wang, M. Zheng, Q. Ruan, Y. Zhou, Y. Chen, P. Dai, Z. Yang, Z. Lin, Y. Long, Y. Li, N. Liu, C.-W. Qiu, J. K. W. Yang, and H. Duan, "Stepwise-nanocavity-assisted transmissive color filter array microprints," *Research* **2018**, 8109054 (2018).
- Y. Hu, X. Luo, Y. Chen, Q. Liu, X. Li, Y. Wang, N. Liu, and H. Duan, "3D-integrated metasurfaces for full-colour holography," *Light Sci. Appl.* **8**, 86 (2019).
- X. Luo, Y. Hu, X. Li, Y. Jiang, Y. Wang, P. Dai, Q. Liu, Z. Shu, and H. Duan, "Integrated metasurfaces with microprints and helicity-multiplexed holograms for real-time optical encryption," *Adv. Opt. Mater.* **8**, 1902020 (2020).
- Y. Hu, L. Li, Y. Wang, M. Meng, L. Jin, X. Luo, Y. Chen, X. Li, S. Xiao, H. Wang, Y. Luo, C. W. Qiu, and H. Duan, "Trichromatic and tripolarization-channel holography with noninterleaved dielectric metasurface," *Nano Lett.* **20**, 994–1002 (2020).
- K.-T. Lee, J. Y. Lee, S. Seo, and L. J. Guo, "Colored ultrathin hybrid photovoltaics with high quantum efficiency," *Light Sci. Appl.* **3**, e215 (2014).
- M. Song, D. Wang, S. Peana, S. Choudhury, P. Nyga, Z. A. Kudyshev, H. Yu, A. Boltasseva, V. M. Shalaev, and A. V. Kildishev, "Colors with plasmonic nanostructures: a full-spectrum review," *Appl. Phys. Rev.* **6**, 041308 (2019).
- K. Kumar, H. Duan, R. S. Hegde, S. C. W. Koh, J. N. Wei, and J. K. W. Yang, "Printing colour at the optical diffraction limit," *Nat. Nanotechnol.* **7**, 557–561 (2012).
- M. L. Tseng, J. Yang, M. Semmlinger, C. Zhang, P. Nordlander, and N. J. Halas, "Two-dimensional active tuning of an aluminum plasmonic array for full-spectrum response," *Nano Lett.* **17**, 6034–6039 (2017).
- B. Yang, W. Liu, Z. Li, H. Cheng, S. Chen, and J. Tian, "Polarization-sensitive structural colors with hue-and-saturation tuning based on all-dielectric nanopixels," *Adv. Opt. Mater.* **6**, 1701009 (2018).
- V. Flauraud, M. Reyes, R. Paniagua-Domínguez, A. I. Kuznetsov, and J. Brugger, "Silicon nanostructures for bright field full color prints," *ACS Photon.* **4**, 1913–1919 (2017).
- H. Shin, M. F. Yanik, S. Fan, R. Zia, and M. L. Brongersma, "Omnidirectional resonance in a metal–dielectric–metal geometry," *Appl. Phys. Lett.* **84**, 4421–4423 (2004).
- D. Zhao, L. Meng, H. Gong, X. Chen, Y. Chen, M. Yan, Q. Li, and M. Qiu, "Ultra-narrow-band light dissipation by a stack of lamellar silver and alumina," *Appl. Phys. Lett.* **104**, 221107 (2014).
- M. A. Kats, R. Blanchard, P. Genevet, and F. Capasso, "Nanometre optical coatings based on strong interference effects in highly absorbing media," *Nat. Mater.* **12**, 20–24 (2013).
- C. Ji, C. Yang, W. Shen, K.-T. Lee, Y. Zhang, X. Liu, and L. J. Guo, "Decorative near-infrared transmission filters featuring high-efficiency and angular-insensitivity employing 1D photonic crystals," *Nano Res.* **12**, 543–548 (2019).
- Z. Li, S. Butun, and K. Aydin, "Large-area, lithography-free super absorbers and color filters at visible frequencies using ultrathin metallic films," *ACS Photon.* **2**, 183–188 (2015).
- Z. Yang, Y. Chen, Y. Zhou, Y. Wang, P. Dai, X. Zhu, and H. Duan, "Microscopic interference full-color printing using grayscale-patterned Fabry-Perot resonance cavities," *Adv. Opt. Mater.* **5**, 1700029 (2017).
- Z. Yang, C. Ji, D. Liu, and L. J. Guo, "Enhancing the purity of reflective structural colors with ultrathin bilayer media as effective ideal absorbers," *Adv. Opt. Mater.* **7**, 1900739 (2019).
- S. So, T. Badloe, J. Noh, J. Bravo-Abad, and J. Rho, "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **9**, 1041–1057 (2020).
- R. S. Hegde, "Deep learning: a new tool for photonic nanostructure design," *Nanoscale Adv.* **2**, 1007–1023 (2020).
- P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photon. Res.* **9**, B182–B200 (2021).
- Q. Zhang, H. Yu, M. Barbiero, B. Wang, and M. Gu, "Artificial neural networks enabled by nanophotonics," *Light Sci. Appl.* **8**, 42 (2019).
- Y. Kiarashinejad, S. Abdollahramezani, and A. Adibi, "Deep learning approach based on dimensionality reduction for designing electromagnetic nanostructures," *npj Comput. Mater.* **6**, 12 (2020).
- S. So, J. Mun, and J. Rho, "Simultaneous inverse design of materials and structures via deep learning: demonstration of dipole resonance engineering using core–shell nanoparticles," *ACS Appl. Mater. Interfaces* **11**, 24264–24268 (2019).

27. J. Jiang and J. A. Fan, "Global optimization of dielectric metasurfaces using a physics-driven neural network," *Nano Lett.* **19**, 5366–5372 (2019).
28. Y. Chen, J. Zhu, Y. Xie, N. Feng, and Q. H. Liu, "Smart inverse design of graphene-based photonic metamaterials by an adaptive artificial neural network," *Nanoscale* **11**, 9749–9755 (2019).
29. T. Zhang, J. Wang, Q. Liu, J. Zhou, J. Dai, X. Han, Y. Zhou, and K. Xu, "Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks," *Photon. Res.* **7**, 368–380 (2019).
30. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. G. DeLacy, J. D. Joannopoulos, M. Tegmark, and M. Soljačić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
31. I. Malkiel, M. Mrejen, A. Nagler, U. Arieli, L. Wolf, and H. Suchowski, "Plasmonic nanostructure design and characterization via deep learning," *Light Sci. Appl.* **7**, 60 (2018).
32. O. Hemmatyar, S. Abdollahramezani, Y. Kiarashinejad, M. Zandehshahvar, and A. Adibi, "Full color generation with Fano-type resonant  $\text{HfO}_2$  nanopillars designed by a deep-learning approach," *Nanoscale* **11**, 21266–21274 (2019).
33. Z. Huang, X. Liu, and J. Zang, "The inverse design of structural color using machine learning," *Nanoscale* **11**, 21748–21758 (2019).
34. J. Baxter, A. Calà Lesina, J.-M. Guay, A. Weck, P. Berini, and L. Ramunno, "Plasmonic colours predicted by deep learning," *Sci. Rep.* **9**, 8074 (2019).
35. L. Gao, X. Li, D. Liu, L. Wang, and Z. Yu, "A bidirectional deep neural network for accurate silicon color design," *Adv. Mater.* **31**, 1905467 (2019).
36. P. R. Wiecha, A. Lecestre, N. Mallet, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nat. Nanotechnol.* **14**, 237–244 (2019).
37. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
38. L. Xu, M. Rahmani, Y. Ma, D. A. Smirnova, K. Z. Kamali, F. Deng, Y. K. Chiang, L. Huang, H. Zhang, S. Gould, D. N. Neshev, and A. E. Miroshnichenko, "Enhanced light-matter interactions in dielectric nanostructures via machine-learning approach," *Adv. Photon.* **2**, 026003 (2020).
39. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
40. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
41. I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning* (MIT Press, 2016).
42. F. W. Billmeyer, *Color Science: Concepts and Methods, Quantitative Data and Formulae* (Wiley, 1983).
43. C. Williams, G. S. D. Gordon, T. D. Wilkinson, and S. E. Bohndiek, "Grayscale-to-color: scalable fabrication of custom multispectral filter arrays," *ACS Photon.* **6**, 3132–3141 (2019).
44. Y. Horie, S. Han, J.-Y. Lee, J. Kim, Y. Kim, A. Arbabi, C. Shin, L. Shi, E. Arbabi, S. M. Kamali, H.-S. Lee, S. Hwang, and A. Faraon, "Visible wavelength color filters using dielectric subwavelength gratings for backside-illuminated CMOS image sensor technologies," *Nano Lett.* **17**, 3159–3164 (2017).
45. Q. Chen, X. Hu, L. Wen, Y. Yu, and D. R. S. Cumming, "Nanophotonic image sensors," *Small* **12**, 4922–4935 (2016).
46. M. Habekost, "Which color differencing equation should be used?" *Int. Circ. Graph. Educ. Res.* **6**, 1–33 (2013).
47. W. S. Mokrzycki and M. Tatol, "Color difference delta E—a survey," *Mach. Graph. Vis.* **20**, 383–411 (2011).
48. A. Slowik and H. Kwasnicka, "Evolutionary algorithms and their applications to engineering problems," *Neural Comput. Appl.* **32**, 12363–12379 (2020).
49. K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: surpassing human-level performance on imagenet classification," in *IEEE International Conference on Computer Vision* (2015), Vol. **2015**, pp. 1026–1034.

# PHOTONICS Research

## Genetic-algorithm-based deep neural networks for highly efficient photonic device design

YANGMING REN,<sup>1,2,†</sup> LINGXUAN ZHANG,<sup>1,2,†</sup> WEIQIANG WANG,<sup>1,2</sup>  XINYU WANG,<sup>1,2</sup> YUFANG LEI,<sup>1,2</sup> YULONG XUE,<sup>1,2</sup> XIAOCHEN SUN,<sup>1,2,3</sup> AND WENFU ZHANG<sup>1,2,4</sup>

<sup>1</sup>State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100049, China

<sup>3</sup>e-mail: sunxiaochen@opt.ac.cn

<sup>4</sup>e-mail: wfuzhang@opt.ac.cn

Received 1 December 2020; revised 26 January 2021; accepted 25 March 2021; posted 30 March 2021 (Doc. ID 416294); published 24 May 2021

While deep learning has demonstrated tremendous potential for photonic device design, it often demands a large amount of labeled data to train these deep neural network models. Preparing these data requires high-resolution numerical simulations or experimental measurements and cost significant, if not prohibitive, time and resources. In this work, we present a highly efficient inverse design method that combines deep neural networks with a genetic algorithm to optimize the geometry of photonic devices in the polar coordinate system. The method requires significantly less training data compared with previous inverse design methods. We implement this method to design several ultra-compact silicon photonics devices with challenging properties including power splitters with uncommon splitting ratios, a TE mode converter, and a broadband power splitter. These devices are free of the features beyond the capability of photolithography and generally in compliance with silicon photonics fabrication design rules. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416294>

### 1. INTRODUCTION

The advent of photonic integrated circuits (PICs) is expected to revolutionize many industries such as optical interconnects [1–3], optical sensing [3–6], and optical computing [7,8]. The progress of PICs demands compact photonic devices with a variety of properties to increasingly integrate more functionality on a chip. The property of a photonic device is highly determined by its geometry, which is usually designed and optimized by an experienced photonic designer based on intuition and numerical simulations. This traditional method is limited by designers' experience and usually only applies to the designs with simple geometric structures. However, complex geometries are often required for realizing ultra-compact photonic devices with challenging properties. The design of these complex structures with a large number of parameters is beyond the capability of human intuition and demands more efficient methods. The brute force optimization method is impractical, as the numerical simulation algorithms, such as the finite-difference time-domain (FDTD) method, are very computation intensive and time consuming. Traditional regression-based optimization methods face difficulties in convergence due to the high-dimensional parameter space. A variety of inverse design methods such as level set methods, adjoint method, local-optimization techniques, and direct-binary

search [9–19] have been developed to tackle this issue. These methods expand the design space and realize many photonic devices with complex and nonintuitive structures that present extraordinary properties. Fundamentally, these algorithms are rule-based approaches containing iterative searching steps in a case-by-case manner, often relying on numerical simulations in each step to produce intermediate results that help to modify the searching strategy. Such stochastic algorithms are limited by their random-search nature and are insufficient for complex designs in a multi-constrained problem [20–22].

This challenge may be addressed by deep learning (DL) and more specifically deep neural networks (DNNs), which use many processing layers to learn the representations of data with multiple levels of abstraction [23]. DNN has been revolutionizing many fields, including image analysis [24], natural language processing [25], materials science [26], quantum physics [27], and so on. Neural network (NN)-based methods have previously been adopted in studying photonic devices such as phase delay structures [28], photonic crystal nanocavities [29], and nanophotonic particles [30]. However, applying a DNN to photonic device design faces a practical challenge of preparing a large training dataset [20]. For example, Ma *et al.* used an NN model trained on 30,000 device designs for a chiral metamaterial inverse design [31]; Tahersima *et al.* used an NN

model trained on 20,000 designs to realize a power splitter with specific splitting ratios [32]. These large amounts of design data are generated by high-resolution 3D FDTD simulations and may take days to complete on a typical workstation.

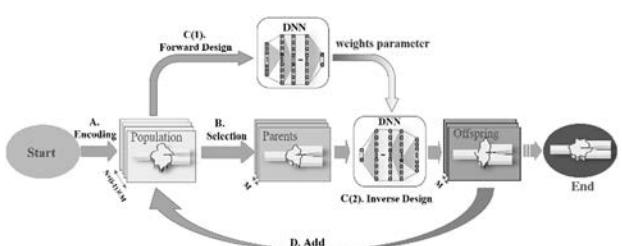
In this paper, we present a genetic-algorithm-based deep neural network (GDNN) method that requires an order of magnitude less simulation data (less than 3000 in the presented cases) for training. By using this approach, we have designed several silicon photonics devices including power splitters with uncommon splitting ratios, a TE mode converter, and a broadband power splitter. Besides high efficiency, the algorithm exhibits great flexibility and ability in dealing with multiple design constraints. For example, the broadband power splitter is designed for as many as 11 optimization targets, and the optimal design exhibits a maximum insertion loss of 0.47 dB and good flatness over 400 nm wavelength range with a  $2.5\text{ }\mu\text{m} \times 1.5\text{ }\mu\text{m}$  footprint. In addition, these devices are free of the features beyond the capability of photolithography and generally in compliance with silicon photonics fabrication design rules.

## 2. ALGORITHM

The GDNN-based inverse design algorithm developed in this work includes encoding, selection, and DNN-based forward and inverse design processes. The workflow of the algorithm is exhibited in Fig. 1. First, the device models and numerical simulation results of the initial device designs are encoded into a dataset called population (step A of Fig. 1). In mimicking the natural selection process, the population is selected with respect to an objective function (step B of Fig. 1). The whole population is also used to generate the weight parameters in the DNN models of both the forward design and the inverse design [step C (1) of Fig. 1]. The selection process produces an adaptive (good performance) generation as the parents to reproduce their offspring through the inverse design process [step C (2) of Fig. 1]. These offspring are then added to the initial population (step D of Fig. 1) and update the weights in the DNN models to start the next iteration. This process is repeated until the DNN model meets the required optimization target to the investigated design problem.

### A. Encoding

The encoding process aims to map a physical problem, e.g., the output optical power of an optical device with multiple degrees of freedom in its physical structures, to a mathematical data structure. An encoded data structure represents a solution to the problem and affects the optimization process. In this work,

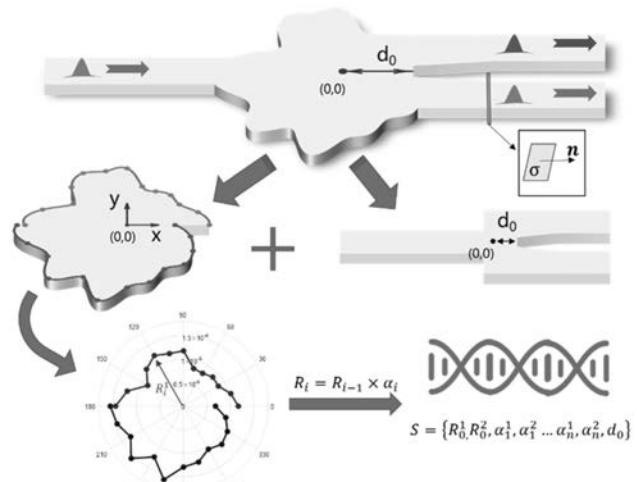


**Fig. 1.** Workflow of the GDNN algorithm developed in this paper.

each individual of the initial population is described with two sets of parameters. One set represents the encoded device physical structure that will be genetically optimized, and the other represents the investigated optical performance. The encoding process of a power splitter device is schematically shown in Fig. 2. An arbitrarily shaped one-by-two power splitter includes one input waveguide, two output waveguides, and a core that is described by its border line. By considering the device design rules and its symmetry, we adopt the polar coordinate system with its origin located on the center line and use two sets of polar vectors ( $R_i^1 = \{R_0^1, R_1^1, \dots, R_n^1\}, R_i^2 = \{R_0^2, R_1^2, \dots, R_n^2\}$ ) to discretize the top (within  $[0, \pi]$ ) and the bottom (within  $[\pi, 2\pi]$ ) part of the border line, respectively. The complete device structure is formed by smoothly connecting these two sets of polar vectors and combining the input and the two output ports. Practically, a foundry generally imposes design rules to disallow very small or sharp structures in a layout. Thus, we also introduce a set of gradient constraints  $\alpha$  ( $S = \{R_0^1, R_0^2, \alpha_1^1, \alpha_1^2, \dots, \alpha_n^1, \alpha_n^2, d_0\}$ ) between the radii of the two adjacent polar vectors. With this encoding process, an individual device is completely described by a design parameter vector  $S = \{R_0^1, R_0^2, \alpha_1^1, \alpha_1^2, \dots, \alpha_n^1, \alpha_n^2, d_0\}$ , where  $d_0$  is the distance of the output branch from the origin as shown in Fig. 2. The investigated performance or the optimization objective is represented by a figure of merit (FOM), which is defined as the optical power in the zeroth transverse electric (TE0) mode of an output waveguide. If the transverse magnetic (TM) component is negligible, which is validated by simulation results, the FOM can be calculated by

$$\text{FOM} = \frac{1}{4} \frac{\int_{\sigma} (E \times H_0^* + E_0^* \times H) \cdot d\sigma}{\int_{\sigma} \text{Re}(E_0 \times H_0^*) \cdot d\sigma}, \quad (1)$$

where  $\sigma$  is the cross section of the waveguide,  $E_0$  ( $H_0$ ) is the electric (magnetic) field distribution of the TE0 mode in the waveguide, and  $E$  ( $H$ ) is the output electric (magnetic) field distribution from the device FDTD simulation. We believe



**Fig. 2.** Encoding process that uses polar vectors and design rule constraints as a parameter vector to describe the design of a given photonic device.

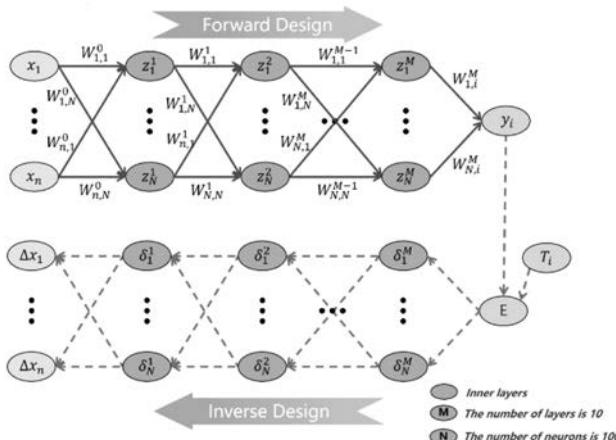
that using the polar coordinate system with carefully designed constraints can effectively describe a sufficiently flexible device geometry while avoiding the violation of foundry design rules. This encoding approach simplifies the complexity of the GDNN inverse design algorithm and can be generalized for other optical device structures beyond the several examples shown in this work.

### B. Selection and DNN-Based Evolution

The population selection process is a prerequisite for DNN-based inverse design. It evaluates the FOM of each individual in the current population and selects a subset of the population with good values to be the parents that “reproduce” the next generation of device designs. Due to the “genetic” correlation between the parents and the offspring connected by the neural networks, additional factors can be added to the selection process to intentionally guide the direction of certain features in the evolution. For example, a punishment factor on the device area can be added to guide the algorithm to favor more compact designs.

A genetic algorithm mainly includes three operations: selection, crossover, and mutation. Crossover and mutation are the key steps to produce excellent offspring. In the GDNN algorithm, we use the DNN model to replace the crossover and mutation operations of the genetic algorithm. The DNN model can be inferred in both forward and inverse ways. The forward design process can obtain the device FOM and network weight parameters in the form of a given encoded design parameter vector. The network weight parameters are the key settings for enabling the reverse design process. The inverse design process after obtaining the network weight parameters utilizes the FOM data associated with wavelength as the input to produce offspring device designs that are correlated but different than the parent. The forward and the inverse design processes are schematically shown in Fig. 3. A fully connected multilayer DNN is used with the sigmoid activation function and back-propagation method for weight updates.

In the inverse design process, the cost function is represented by the error  $E$  of the FOM of a specific design ( $S = \{x_1, \dots, x_n\}$ ,  $Y = \{\text{FOM}_1, \text{FOM}_2\}$ ) from its ideal value  $T$ ,



**Fig. 3.** Schematic drawing of the DNN models of the forward and inverse design processes.

i.e.,  $E = \sum_{i=1}^m (\text{FOM}_i - T_i)^2$ . This error is back-propagated by the equation

$$\delta_i^k = \frac{\partial E}{\partial Z_i^k} = \sum_{j=1}^N \left( \frac{\partial E}{\partial Z_j^{k+1}} \cdot \frac{\partial Z_j^{k+1}}{\partial Z_i^k} \right) = \sum_{j=1}^N \left( \delta_j^{k+1} \cdot \frac{\partial Z_j^{k+1}}{\partial Z_i^k} \right), \quad (2)$$

where  $Z_i^k$  is the weight of the neurons in the  $k$ th layer, i.e.,  $Z_i^k = \sum_{j=1}^N f(W_{j,i}^{k-1} + Z_j^{k-1})$ , where  $f$  is the activation function. By repeating this step, the error eventually propagated to the input layer gives the offspring device design variation

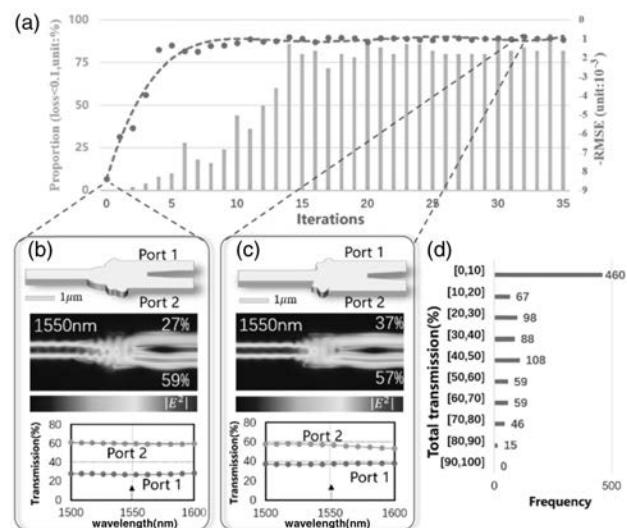
$$\Delta x_i = \sum_{j=1}^N \left( \delta_j^1 \cdot \frac{\partial Z_j^1}{\partial x_j} \right) = \sum_{j=1}^N [\delta_j^1 \cdot f'(x_j)], \quad (3)$$

where  $f'(x_j)$  is the derivative of the activation function.

### 3. SIMULATION RESULTS

In order to study its capability and performance, we applied the GDNN algorithm to design several very compact Si photonics devices as examples, including low-loss power splitters with specific unequal output powers, a TE mode converter, and a broadband power splitter. All of these devices are designed based on common Si photonics platforms with silicon-on-insulator (SOI) substrates with 220 nm device silicon and a 2–3  $\mu\text{m}$  buried oxide (BOX) layer. The waveguide is 500 nm in width and operated at (or around) 1550 nm wavelength. All the devices are designed according to the process design rule that no line or space feature size is smaller than 0.2  $\mu\text{m}$ .

We take a power splitter with splitting ratio of 2:3 as an example to quantitatively analyze the design process. The initial population is composed of 1000 device structures randomly generated according to the encoding process stated earlier. The distribution of total optical transmission of a typical random population is shown in Fig. 4(d). In this case, the majority



**Fig. 4.** Design analyses of a power splitter with splitting ratio of 2:3: (a) the evolution of the qualified population proportion; (b) and (c) the FDTD simulation result of the best devices in the initial population and the final population; (d) the distribution of optical transmission of the initial population.

of the population show very poor transmission, and only 1.5% transmit >80% light. The FDTD simulation result of the best device in this initial population is shown in Fig. 4(b). Then the DNN-based inverse and forward design processes are iterated with 50 offspring designs produced in each cycle. The evolution of the population proportion that exceeds a certain optical performance target, e.g., transmission >90%, with respect to the number of design iterations (generations) is shown in Fig. 4(a). The evolution plot shows the growth of performance with iterations until it saturates to about 80%. Another way to study the convergence is the cost function mentioned earlier, i.e., the root-mean-square error (RMSE) of the FOM with respect to a target value. The RMSE as a function of generation is also shown in Fig. 4(a) and clearly exhibits a convergence behavior. The algorithm generates a robust high-percentile good performance as well as low variations beyond about 14 generations. The FDTD simulation result of the final device design generated from this algorithm is shown in Fig. 4(c). The device has a size of only  $2\text{ }\mu\text{m} \times 1.7\text{ }\mu\text{m}$ , 0.27 dB insertion loss, and 2:00:3.08 splitting ratio.

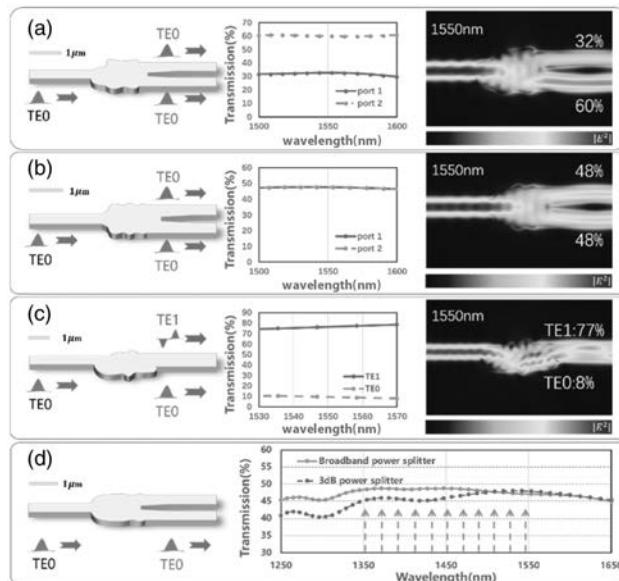
Similar design processes are adopted to optimize the other optical devices including a 1:2 power splitter with 0.36 dB insertion loss and  $2.2\text{ }\mu\text{m} \times 1.8\text{ }\mu\text{m}$  footprint [Fig. 5(a)], a 1:1 power splitter with 0.18 dB insertion loss and  $2\text{ }\mu\text{m} \times 1.5\text{ }\mu\text{m}$  footprint [Fig. 5(b)], a TE mode converter with 77% TE0-to-TE1 mode conversion efficiency and  $3\text{ }\mu\text{m} \times 1.5\text{ }\mu\text{m}$  footprint [Fig. 5(c)], and a broadband power splitter operating over a 400 nm range [Fig. 5(d)]. The TE mode converter is an order of magnitude smaller than a conventional directional coupler-based design [33,34]. The successful design of the broadband power splitter shows the capability of multi-objective optimization in our algorithm. The broadband power splitter is optimized from the 1:1 power splitter shown earlier by introducing optimization objectives at 11 wave-

lengths evenly distributed between 1350 and 1550 nm. The multi-objective GDNN model optimizes the power splitter design [Fig. 5(d)] with a maximum insertion loss of 0.47 dB and a maximum loss variation of 0.36 dB within 400 nm (1250–1650 nm) and  $2.5\text{ }\mu\text{m} \times 1.5\text{ }\mu\text{m}$  footprint. More importantly, the maximum insertion loss is 0.2 dB and the maximum loss variation is 0.09 dB within the 200 nm (1350–1550 nm) design range. In comparison, the 1:1 power splitter shown in Fig. 5(b) shows more than double insertion loss and loss variation. The designed broadband power splitter presents good performance from the O band to the C band and is useful in applications demanding wide-band operations.

#### 4. DISCUSSIONS AND CONCLUSION

In this work, the DNN was implemented by using PyTorch, an open source machine learning framework. The same GDNN architecture is used for all the designs. The architecture includes 10 hidden layers with the dimension of 100. There are 1000 trained data for the initial population and 50 reproduced offspring for each iteration. The algorithm is implemented on a workstation with an AMD Ryzen 3700X CPU and a Nvidia GeForce RTX 2070 GPU. It takes about 6–10 h to complete the whole design including the FDTD simulation of the 1000 initial population data and the whole genetic optimization process.

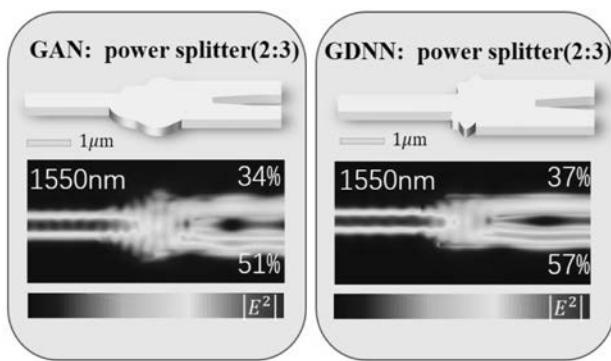
DNN has been proved to be a powerful tool for designing optical devices using an inverse design process. Previous DNN-based inverse design algorithms usually require 20,000 to 40,000 training data [31,32], which take the majority of the design time due to the computation intensive high-resolution FDTD simulations. The use of genetic optimization reduces the number of the training data by an order of magnitude. In addition, unlike many inverse design algorithms, our GDNN algorithm does not require the initial data set to include a large amount of good performance designs because the quality of the population is iteratively improved during the evolution. In fact, we intentionally include some device designs, violating design rules that serve for regulating the constraints of the inverse design. As shown in Table 1, the GDNN uses 1000 sets of data for the initial training of the model. After that, 50 sets of offspring will be added to participate in the training of the model. It is worth mentioning that the model design with similar structural design uses the same set of initial datasets (1000 sets of



**Fig. 5.** GDNN design examples with transmission spectrum and FDTD simulation results: (a) a 1:2 power splitter, (b) a 1:1 power splitter, (c) a TE mode converter, and (d) a broadband power splitter.

**Table 1. Training Data Summary of the Designs in This Work**

Device Designs	Initial Data	Number of Iterations	Number of Offspring	Total Data
Power splitter (1:1)	1000	35	50	2750
Power splitter (1:2)	1000	28	50	2400
Power splitter (2:3)	1000	32	50	2600
TE mode converter	1000	30	50	2500
Broadband splitter	1000	23	50	2150



**Fig. 6.** Comparison of GAN and GDNN design results.

data), for instance, power splitter (1:1) and broadband splitter, power splitter (1:2) and power splitter (2:3). Our GDNN algorithm completes a design task with less than 3000 design data including the 1000 initial population and the offspring of all iterations. This is an order of magnitude less than the training data for a typical previous inverse design algorithm.

In comparison, we have investigated using an efficiency-focused generative adversarial network (GAN) model to design the same photonic devices (2:3 power splitter). The results (Fig. 6) show that greater than 10,000 device design training data are required to achieve designs with comparable performance.

In conclusion, by combining DNN with genetic evolution, our GDNN method exhibits great flexibility and efficiency in designing ultra-compact photonic devices with challenging properties. Compared with previous optimization algorithms, the GDNN algorithm is more capable of complex multi-constraint and multi-objective optimization. Without relying on pixelized features in a predefined geometry and grid, our method allows design of more flexible device structures while obeying fabrication design rules. The genetic evolution nature of the algorithm greatly increases computation efficiency and reduces FDTD training data by an order of magnitude compared with previous DNN-based inverse design methods. In addition, the genetic selection process allows introduction of biases to guide certain design directions such as a small footprint. To exhibit the capability of the method, we design several Si photonics devices including power splitters with uncommon splitting ratios, a TE mode converter, and a broadband power splitter. These devices are free of the features beyond the capability of photolithography and generally in compliance with silicon photonics fabrication design rules. We believe the presented GDNN algorithm can be widely implemented in designing many complex micro- and nanophotonic structures that could not be easily realized before.

**Funding.** Strategic Priority Research Program of the Chinese Academy of Sciences (XDB24030600); National Natural Science Foundation of China (12004421, 61635013, 61675231); Youth Innovation Promotion Association of Chinese Academy of Sciences (2016535); West Light Foundation of the Chinese Academy of Sciences (XAB2017A09); Natural Science Basic Research Program of

Shaanxi (2019JQ-447); Research Project of Xi'an Postdoctoral Innovation Base (201903).

**Acknowledgment.** The authors thank Dr. Shiyue Hua for helpful discussion.

**Disclosures.** The authors declare that they have no conflicts of interest.

<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

- M. J. R. Heck, H. W. Chen, A. W. Fang, B. R. Koch, D. Liang, H. Park, M. N. Sysak, and J. E. Bowers, "Hybrid silicon photonics for optical interconnects," *IEEE J. Sel. Top. Quantum Electron.* **17**, 333–346 (2011).
- Y. Urino, T. Usuki, J. Fujikata, M. Ishizaka, K. Yamada, T. Horikawa, T. Nakamura, and Y. Arakawa, "High-density and wide-bandwidth optical interconnects with silicon optical interposers [Invited]," *Photon. Res.* **2**, A1–A7 (2014).
- Y. Li, Y. Zhang, L. Zhang, and A. W. Poon, "Silicon and hybrid silicon photonic devices for intra-datacenter applications: state of the art and perspectives," *Photon. Res.* **3**, B10–B27 (2015).
- J. Sun, E. Timurdogan, A. Yaacobi, E. S. Hosseini, and M. R. Watts, "Large-scale nanophotonic phased array," *Nature* **493**, 195–199 (2013).
- X. Sun, L. Zhang, Q. Zhang, and W. Zhang, "Si photonics for practical LiDAR solutions," *Appl. Sci.* **9**, 4225 (2019).
- C. V. Poult, M. J. Byrd, M. Raval, Z. Su, N. Li, E. Timurdogan, D. Coolbaugh, D. Vermeulen, and M. R. Watts, "Large-scale silicon nitride nanophotonic phased arrays at infrared and visible wavelengths," *Opt. Lett.* **42**, 21–24 (2017).
- J. Feldmann, N. Youngblood, C. D. Wright, H. Bhaskaran, and W. H. P. Pernice, "All-optical spiking neurosynaptic networks with self-learning capabilities," *Nature* **569**, 208–214 (2019).
- K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," *Nat. Commun.* **5**, 3541 (2014).
- K. Xu, L. Liu, X. Wen, W. Sun, N. Zhang, N. Yi, S. Sun, S. Xiao, and Q. Song, "Integrated photonic power divider with arbitrary power ratios," *Opt. Lett.* **42**, 855–858 (2017).
- E. Khoram, X. Qian, M. Yuan, and Z. Yu, "Controlling the minimal feature sizes in adjoint optimization of nanophotonic devices using b-spline surfaces," *Opt. Express* **28**, 7060–7069 (2020).
- J. Lu and J. Vučković, "Nanophotonic computational design," *Opt. Express* **21**, 13351–13367 (2013).
- C. M. Lalau-Keraly, S. Bhargava, O. D. Miller, and E. Yablonovitch, "Adjoint shape optimization applied to electromagnetic design," *Opt. Express* **21**, 21693–21701 (2013).
- B. Shen, P. Wang, R. Polson, and R. Menon, "An integrated-nanophotonics polarization beamsplitter with  $2.4 \times 2.4 \mu\text{m}^2$  footprint," *Nat. Photonics* **9**, 378–382 (2015).
- D. Vercruyse, N. V. Sapra, L. Su, R. Trivedi, and J. Vučković, "Analytical level set fabrication constraints for inverse design," *Sci. Rep.* **9**, 8999 (2019).
- A. Y. Piggott, E. Y. Ma, L. Su, G. H. Ahn, N. V. Sapra, D. Vercruyse, A. M. Netherton, A. S. P. Khope, J. E. Bowers, and J. Vučković, "Inverse-designed photonics for semiconductor foundries," *ACS Photon.* **7**, 569–575 (2020).
- A. Y. Piggott, J. Lu, K. G. Lagoudakis, J. Petykiewicz, T. M. Babinec, and J. Vucković, "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," *Nat. Photonics* **9**, 374–377 (2015).
- C. Y. Kao, S. Osher, and E. Yablonovitch, "Maximizing band gaps in two-dimensional photonic crystals by using level set methods," *Appl. Phys. B* **81**, 235–244 (2005).

18. K. Wang, X. Ren, W. Chang, L. Lu, D. Liu, and M. Zhang, "Inverse design of digital nanophotonic devices using the adjoint method," *Photon. Res.* **8**, 528–533 (2020).
19. J. C. C. Mak, C. Sideris, J. Jeong, A. Hajimiri, and J. K. S. Poon, "Binary particle swarm optimized 2 × 2 power splitters in a standard foundry silicon photonic platform," *Opt. Lett.* **41**, 3868–3871 (2016).
20. W. Ma, Z. Liu, Z. A. Kudyshev, A. Boltasseva, W. Cai, and Y. Liu, "Deep learning for the design of photonic structures," *Nat. Photonics* **15**, 77–90 (2020).
21. W. Li, F. Meng, Y. Chen, Y. Fan Li, and X. Huang, "Topology optimization of photonic and photonic crystals and metamaterials: a review," *Adv. Theor. Simul.* **2**, 1900017 (2019).
22. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," *Nat. Photonics* **12**, 659–670 (2018).
23. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
24. B. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image net classification with deep convolutional neural networks," *Commun. ACM* **60**, 84–90 (2012).
25. T. Young, D. Hazarika, S. Poria, and E. Cambria, "Recent trends in deep learning based natural language processing," *IEEE Comput. Intell. Mag.* **13**, 55–75 (2018).
26. Y. Mao, Q. He, and X. Zhao, "Designing complex architected materials with generative adversarial networks," *Sci. Adv.* **6**, eaaz4169 (2020).
27. G. Carleo and M. Troyer, "Solving the quantum many-body problem with artificial Neural networks," *Science* **355**, 602–606 (2007).
28. D. Liu, Y. Tan, E. Khoram, and Z. Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photon.* **5**, 1365–1369 (2018).
29. T. Asano and S. Noda, "Optimization of photonic crystal nanocavities based on deep learning," *Opt. Express* **26**, 32704–32717 (2018).
30. J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. Delacy, M. Tegmark, J. D. Joannopoulos, and M. Soljacić, "Nanophotonic particle simulation and inverse design using artificial neural networks," *Sci. Adv.* **4**, eaar4206 (2018).
31. W. Ma, F. Cheng, and Y. Liu, "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS Nano* **12**, 6326–6334 (2018).
32. M. H. Tahersima, K. Kojima, T. Koike-Akino, D. Jha, B. Wang, C. Lin, and K. Parsons, "Deep neural network inverse design of integrated photonic power splitters," *Sci. Rep.* **9**, 1368 (2019).
33. D. Dai and J. E. Bowers, "Novel concept for ultracompact polarization splitter-rotator based on silicon nanowires," *Opt. Express* **19**, 10940–10949 (2011).
34. Y. Xiong, D.-X. Xu, J. H. Schmid, P. Cheben, S. Janz, and W. N. Ye, "Fabrication tolerant and broadband polarization splitter and rotator based on a taper-etched directional coupler," *Opt. Express* **22**, 17458–17465 (2014).

# PHOTONICS Research

## Toward simple, generalizable neural networks with universal training for low-SWaP hybrid vision

BAURZHAN MUMINOV, ALTAI PERRY, RAKIB HYDER, M. SALMAN ASIF, AND LUAT T. VUONG\* 

University of California Riverside, Riverside, California 92521, USA

\*Corresponding author: LuatV@UCR.edu

Received 7 December 2020; revised 9 March 2021; accepted 25 March 2021; posted 19 April 2021 (Doc. ID 416614); published 14 June 2021

Speed, generalizability, and robustness are fundamental issues for building lightweight computational cameras. Here we demonstrate generalizable image reconstruction with the simplest of hybrid machine vision systems: linear optical preprocessors combined with no-hidden-layer, “small-brain” neural networks. Surprisingly, such simple neural networks are capable of learning the image reconstruction from a range of coded diffraction patterns using two masks. We investigate the possibility of generalized or “universal training” with these small brains. Neural networks trained with sinusoidal or random patterns uniformly distribute errors around a reconstructed image, whereas models trained with a combination of sharp and curved shapes (the phase pattern of optical vortices) reconstruct edges more boldly. We illustrate variable convergence of these simple neural networks and relate learnability of an image to its singular value decomposition entropy of the image. We also provide heuristic experimental results. With thresholding, we achieve robust reconstruction of various disjoint datasets. Our work is favorable for future real-time low size, weight, and power hybrid vision: we reconstruct images on a 15 W laptop CPU with 15,000 frames per second: faster by a factor of 3 than previously reported results and 3 orders of magnitude faster than convolutional neural networks. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.416614>

### 1. INTRODUCTION

Image reconstruction has wide application in medicine [1,2], biology [3], X-ray crystallography [4], and low-light vision, among other technologies. These reconstructions generally involve solving an inverse problem and retrieving the phase from phaseless intensity measurements. The field has been an active area of research for several decades [5–7] and inverse solvers achieve impressive results with additional coded optics or optical scanning [8–21]. More recently, deep neural networks, and specifically convolutional neural networks, enable single feed-forward, noniterative reconstruction [22] and are capable of learning from the statistical information contained in a variety of systems, from speckle [23,24] to coded diffraction [25] patterns. Inverse solvers using neural networks are generally faster than iterative, optimization-based, or optical scanning-based algorithms and may require as few as 100 illumination training patterns, for example, with an “unrolled” neural network [26].

However, despite the benefits of using neural networks to solve inverse problems, there are also drawbacks. Some of these issues—especially those associated with phase retrieval—have been solved. Others—particularly those related to generalizability, robustness, and processing time or energy—remain active areas of research [24]. Since neural networks learn how to weigh the importance of information patterns based on training data,

they exhibit a tendency to “memorize” patterns to gain intuition about the task [27]. This predisposition toward prior data is advantageous for building “inductive, artificial intelligence machines” that extract patterns; however, that predisposition is a detriment to the generalizability of inverse solutions, e.g., for building real-time computational cameras. Antun *et al.* [28] highlight three specific issues encountered by neural nets in imaging tasks:

1. Small, sometimes undetectable perturbations in the input (both image and sampling domain) can cause severe artifacts in the image reconstruction.
2. Small structural changes can be left undetected.
3. More samples in the training set can lead to a deterioration of the results (as a result of the “memory” effect described above). Subsequently, algorithms themselves can stall or experience instabilities.

Whereas biomedical applications are aimed at large-image, high-quality image reconstruction [3], we turn our attention toward building real-time computational cameras for low size, weight, and power (SWaP) image reconstruction, which are needed for autonomous-vehicle applications. In our prior effort [29], we demonstrate reconstruction with a “small brain” dual-layer neural network. Such regression-based approaches [30] demonstrate fast reconstruction rates, robustness to noise, and show potential for generalization with a phase vortex

encoder. Here, we focus entirely on the generalizability of a simple neural network using a single-layer architecture for image reconstruction. We supply the model with a generalized or universal training set (UTS) (synthetic images, used to train the neural network) and then test the neural network with images of different, unseen classes [see Fig. 1(a)]. A UTS-trained model overcomes the challenges associated with the “stereotypes” that generally arise from training by a specific image set. On the other hand, some disadvantages include the fact that the neural network is too simple to reconstruct images when nonlinear transformations are required [31]. Nevertheless, our results provide insight for training generalizable neural networks and computational cameras that operate at fast speeds. Our proposed method can readily be used for the initialization of alternating minimization problems or downstream image analysis tasks [32–34].

It is perhaps surprising that the simple learning model possesses enough capacity to recover a good approximation of the inverse coded-diffraction problem, and even with such a simple neural network there are interesting issues to address. In an effort to move toward producing a generalized training set, we compare the performance of the vortex encoder with other random encoders. From there, we build intuition for the UTS design based on the modal decomposition of the training, diffracted imaging patterns, and singular value decomposition (SVD)-entropy. We also perform experiments, which build heuristics for real-world applications. We find that the choice of training images and optical encoder is important for achieving generalizability, since not all imaged patterns provide a

unique mapping to be learned and not all learned intensity patterns aid image reconstruction. While we have not quantitatively analyzed the image reconstruction, i.e., compared the set of training images to the span of the neural network, we observe that reduced SVD-entropy in the training set increases the learning efficiency, in both simulations and experiments.

## 2. PROJECT SETUP

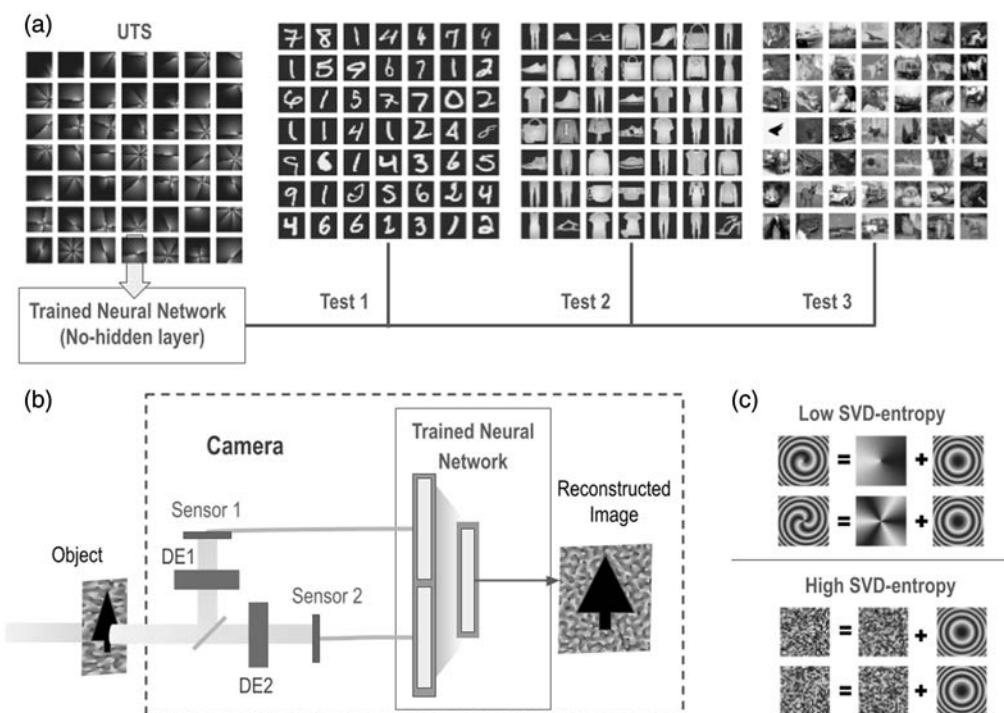
In this section, we review an approach similar to Ref. [29] for our study of generalizable training. Figure 1(b) shows a schematic of the hybrid machine vision system, which encodes the image prior to the neural network with either a random or vortex phase pattern.

### A. Hybrid Vision System

The fields from the object at the diffractive encoder plane are  $F(x, y)$ . The encoder plane is imprinted with two diffractive element patterns  $M(x, y)$ , as shown in Fig. 1(c). A sensor or detector captures the intensity pattern of the electric fields  $F'(u, v)$ . Let  $\mathcal{F}$  be the Fourier transform operation  $(x, y) \rightarrow (u, v)$ , where we capture an image in the Fourier plane:

$$F'(u, v) = \mathcal{F}[M(x, y)F(x, y)]. \quad (1)$$

Light from each object produces two images, each with a different diffractive element  $M(x, y)$ . Although the mask pattern may imprint vector (i.e., polarization-dependent) or spectral (time-dependent) delays, here we assume a homogeneous polarization, a linear encoder, and monochromatic,



**Fig. 1.** (a) Project objective: design a generalized training set for a neural network, which can later be used for general image reconstruction without retraining and can operate in real time. (b) Schematic of hybrid vision camera where light from an object is transmitted through a diffractive encoder (DE). Sensors capture two transmitted images that are combined as inputs to the trained neural network, which reconstruct the object from the detector-plane images. (c) This project employs two pairs of diffractive encoders: one with low SVD-entropy (lens and topological charge  $m = 1$  and 3) and the other with high SVD-entropy (uniformly distributed random pattern).

continuous-wave light. All optical neural networks have been previously demonstrated, notably with several diffractive layers in the terahertz regime [35], with nonlinear activations via saturable-absorbing nonlinearities [36], and with nano-interferometric etalons [37] in the visible regime. All-optical methods maximize speed and minimize energy loss in the neural computation [38]. At the same time, all-optical systems require nonlinear interactions as proxies for the electronic neural network layer activations. These nonlinearities occur at small length scales in order to confine light sufficiently, so all-optical computing may be more sensitive to environmental conditions and less suitable for autonomous-vehicle computational cameras.

By contrast, we focus on hybrid imaging in which optical processing conditions sensor measurements and an electronic neural network performs reconstruction [39–41]. Our work is also inspired by pytchography approaches in Refs. [10–12]. Two phase masks are used to capture the intensity measurements of the object on the sensor, which are then fed to a no-hidden-layer neural network. At this time, we do not predict depth sensing with imaging, so the masks contain lenses for Fourier-plane detection. Here we reproduce the object based on the detector intensity patterns and assume that the detector is in the focal plane associated with a quadratic radial phase of the mask. In recent work, Fresnel mid-field imaging shows potential for better object-based depth detection [42].

In a manner similar to Ref. [29], we generate phase-modulated patterns:

$$F(x, y)M(x, y) = e^{i\alpha X}G(x, y)M(x, y), \quad (2)$$

where  $G(x, y)$  is the Gaussian beam pattern illuminating the object and  $X$  is the positively valued original image. This Gaussian pattern represents a smooth pupil function or the illuminating beam. In our study, we fix  $\alpha = \pi$  and find that the reconstruction quality does not change significantly when  $\alpha$  varies from  $\pi/4$  to  $3\pi/2$ .

The general inverse problem for mapping the detector measurements to the original image involves solving the following nonlinear system of equations:

$$Y = H(X) + N, \quad (3)$$

or for our specific case,

$$Y = |\mathcal{F}[e^{i\alpha X}G(x, y)M(x, y)]|^2 + N, \quad (4)$$

where  $Y$  is the positively valued sensor measurement;  $H(\cdot)$  is a nonlinear transform operator that includes the transfer function of the optics, light scattering, and the sensitivity curve of the detector; and  $N$  is the measurement noise.

The Fourier-plane intensity patterns  $Y$  are the inputs to a neural network. The neural network estimates  $X$  (size  $28 \times 28$ ) given  $Y$  (size  $28 \times 28 \times 2$ ). To train the neural network, we use the TensorFlow library with the mean squared error loss and Adam optimization algorithm. Convergence is achieved with similar results using either “linear” or “ReLU” activation. Our approach is simple and shows promising opportunities for generalized image reconstruction with “small brain” neural networks.

## B. Universal Training Sets and Diffractive Encoders

We choose two pairs of diffractive encoders. One pair is composed of vortex masks, where each mask has an on-axis singularity of either  $m = 1$  or  $3$ :

$$M(x, y) = e^{-(x^2+y^2)\left(\frac{j}{f\lambda}+\frac{1}{w^2}\right)}e^{im\phi}, \quad (5)$$

where  $f$  is the effective focal length of the radial quadratic phase,  $\lambda$  is the wavelength of light,  $m$  is an on-axis topological charge, and  $w$  is the width of the Gaussian beam illuminating the mask. Figures 1(b) and 1(c) show diffractive elements with  $m = 1$  and  $3$ . The second pair is composed of random masks, where each pixel of the transmitted pattern is encoded with a random phase from  $0$  to  $2\pi$ . The mask is also illuminated with the same Gaussian beam. On the side of the training, we work with a range of images composed of  $28 \times 28$  patterns that are random  $X_R$ , Fourier-based  $X_F$ , or shapes related to a vortex phase  $X_V$ .

We approach the generalized training to understand the modal distribution of each image  $X$ . In principle, the training images should span the space of the test images, which defines the requirements for reconstruction. This would suggest that each coded-diffraction Fourier-plane image should be decomposed into Fourier modes, since this common basis provides a unique and straightforward basis for each image. Such Fourier patterns are linear wave patterns that change with phase and vary with variables  $j, k, l, n$ :

$$X_{F(s_j, s_k, \phi_l, n)}(x, y) = \angle[e^{i(xs_j+ys_k+\phi_l)}]G_n, \quad (6)$$

where combinations of  $s_j = 2\pi j/dx, s_k = 2\pi k/dy$ , and  $k$  span the Fourier space intended to reproduce any arbitrary image and  $N$ .  $G_n$  represents a scanning Gaussian beam with varied width and center,

$$G_n(x, y) = e^{-[(x-x_n)^2+(y-y_n)^2]/w_n^2}, \quad (7)$$

where  $x_n, y_n, w_n$  tune size of the UTS to be comparable to others. The size of the dataset also changes the phase shift, where  $\phi_k = 2\pi k/N$  and  $N$  is the number of the uniquely valued wave fringes with wavenumbers  $s_j, s_k$  in  $X_F$ .

We refer to a “vortex training set” as a UTS composed of shapes similar to the phases of a vortex beam that have distinct edges and curves:

$$X_{V(x_j, y_k, \phi_j, n, l)}(x, y) = \angle\{e^{im_l \arctan[(y-y_k)/(x-x_j)]+\phi_k}\}G_{j,k,n}. \quad (8)$$

For the vortex  $X_V$  as well as the random  $X_R$  UTS, we use uniformly distributed random variables to mask the pattern with a Gaussian profile. In other words, combinations of  $x_j, y_j$ , and  $\phi_k = 2\pi k/N$  span the dataset, or

$$G_{j,k,n}(x, y) = e^{-[(x-x_j)^2+(y-y_k)^2]/w_n^2}. \quad (9)$$

This Gaussian function  $G_{j,k,n}(x, y)$  represents a scanning light beam that illuminates the training images. All image patterns are positively valued and normalized to have a peak value of 1.

We produce three UTSs that span the image space using up to 40,000 patterns. The goal of our project is to illustrate trends and intuition with these datasets.

Once trained with a large dataset, we observe that the dense neural network without hidden layers can approximate almost

any shape-based image (MNIST, fashion MNIST, CIFAR). An example set of reconstructed images from different classes is shown in Fig. 2. Figure 2 shows a representative set of images reconstructed from models trained with  $\mathbf{X}_F$ ,  $\mathbf{X}_V$ , and  $\mathbf{X}_R$  and a vortex mask. In each case, 20,000 training images are used. Error with thresholding is as low as 10% with test datasets. While the overall error is similar, models trained with the vortex-phase datasets,  $\mathbf{X}_V$ , generally have the lowest error and strongly highlighted edges. Meanwhile models trained with a Fourier basis  $\mathbf{X}_F$  have the highest error and models trained with a random basis  $\mathbf{X}_R$  have error in between, with error distributed over the area of the image. Additional differences are explained in the following section.

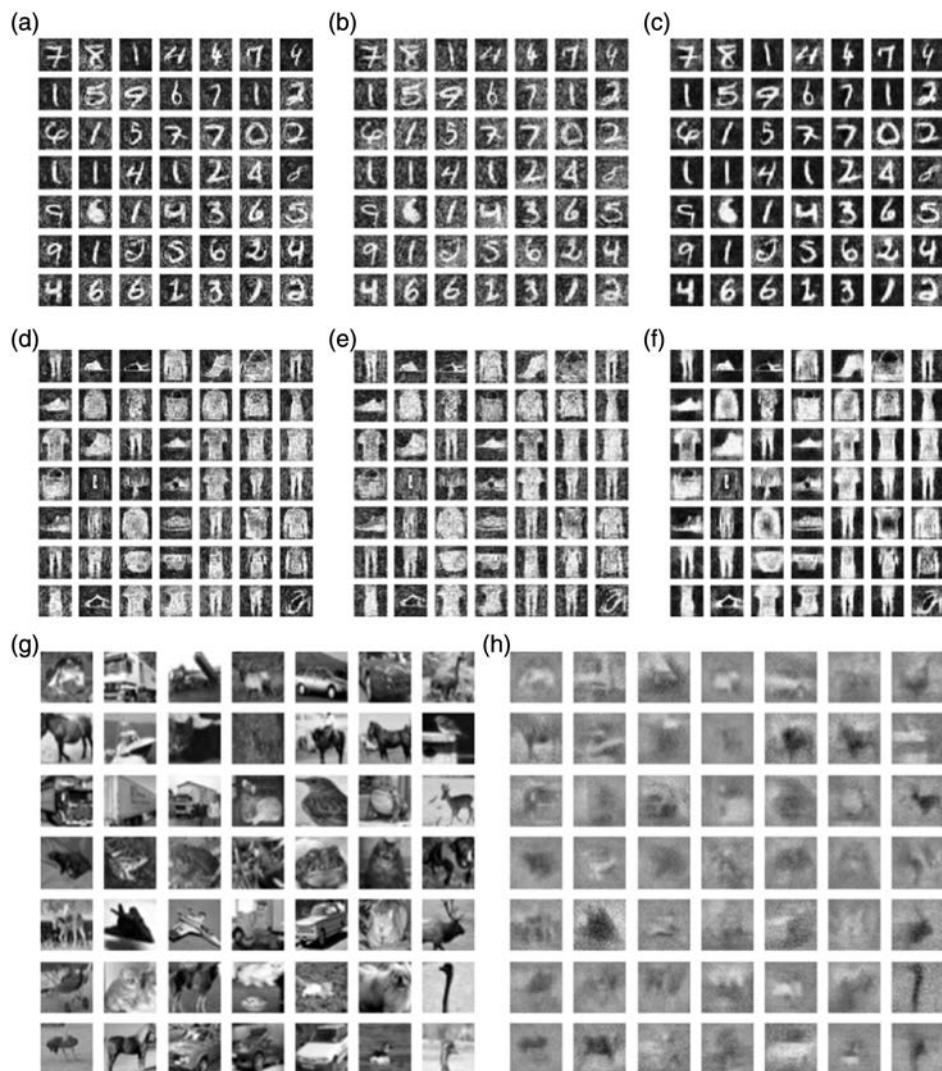
### C. Differences in Convergence and Single-Pixel Response with Different Training Sets

With this simple neural network and three different UTSs, we observe trends in convergence and overfitting. These trends consistently depend on the choice of the UTS patterns regardless

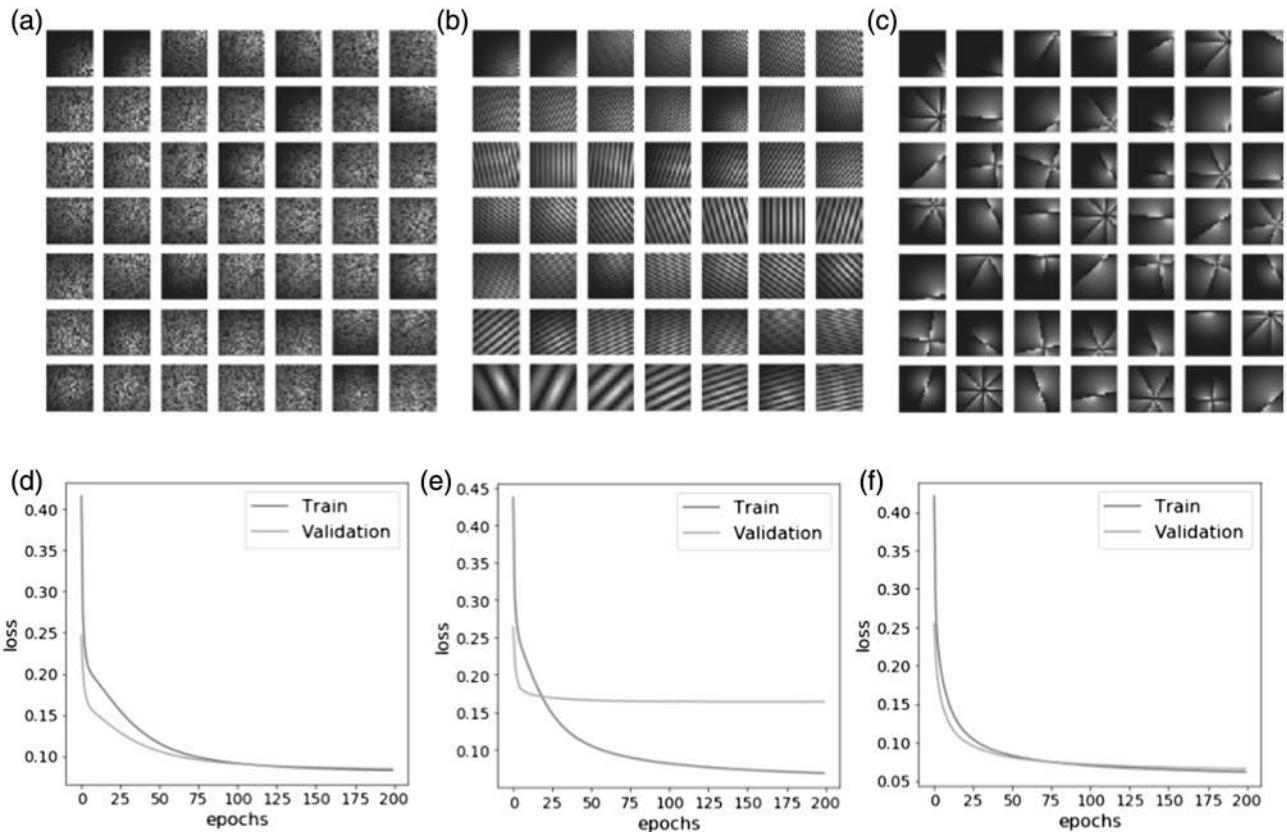
of the choice of mask  $M_V$  or  $M_R$ . Figures 3(a)–3(c) show samples from 20,000-image  $\mathbf{X}_F$ ,  $\mathbf{X}_V$ , and  $\mathbf{X}_R$  UTS with the vortex mask  $M_V$ . Some pairings converge with minimal overfitting while others do not provide enough information in  $\mathbf{Y}$  to calculate the inverse of the nonlinear mapping,  $H(\mathbf{X})$  [Figs. 3(d)–3(f)].

A Fourier basis is the most well-known spectral basis for decomposing an image. When training with a Fourier basis, the validation loss stops decreasing after a certain number of epochs, which signals that the neural network struggles to extract information about the mapping given this orthogonal set of images. What this tells us is rather unintuitive about the span or basis of image reconstruction with neural networks, but potentially addressed in Ref. [43]: the images are less effectively learned by the neural net because there is minimal overlap between them; the correlations between Fourier modes are less visible to the neural net.

The random UTS also unreliable converges when the dataset is smaller than 2000, and its loss generally shows a “hill,”



**Fig. 2.** Reconstructed images from (a), (b), (c) MNIST handwritten and (d), (e), (f) fashion MNIST datasets with random, Fourier, and vortex bases, respectively. The vortex basis provides edge enhancement for object detection. (g) Ground truth and (h) reconstructed images from the CIFAR-10 dataset using the vortex training bases and a vortex mask as the encoder.



**Fig. 3.** (a)–(c) Sample training images  $X_R$ ,  $X_F$ , and  $X_V$  or random, Fourier, and vortex training sets. (d)–(f) Corresponding training and validation curves.

where the loss plateaus before dropping. Meanwhile, the vortex-based UTS is less prone to such behavior. This combination of trends tells us that neither orthogonality nor randomness is ideal for training a neural network. The structured pattern of our vortex-based UTS  $X_V$  is a better candidate for generalized training compared to random  $X_R$  or Fourier  $X_F$  patterns. In our discussion, we provide some measures related to the UTS image analysis and trained model robustness.

### 3. DISCUSSION

In this section, we discuss the ability to recreate sharp images, which may be seen by the single-pixel response. The single-pixel response from the random UTS-trained neural network is sharply corrugated [Fig. 4(a)], whereas the structured, single-pixel images from the vortex-trained model are generally smooth with a sharp “hole” in the center or dark spot [Fig. 4(b)]. We claim that these differences in the impulse response are responsible for the edge-enhanced reconstruction of shapes in Figs. 2(c) and 2(f). Figures 4(a) and 4(b) illustrate example images reconstructed with just one “hot” pixel in the camera sensor plane. These patterns are the building blocks of the reconstruction scheme and these patterns change depending on how the model is trained. Depending on the training set, the model is tuned to pay attention to different features of the image, which may depend on the task at hand.

Figure 4(c) provides a simple noise analysis that shows the additional advantage of robustness when the neural network is trained with a low-entropy UTS. We show the reconstruction error as a function of noise magnitude. Poisson shot noise and background noise are added to the Fourier-plane intensity patterns of the test image set. Low SVD-entropy image training and encoders appear more robust.

#### A. Analysis with Singular Value Decomposition Entropy

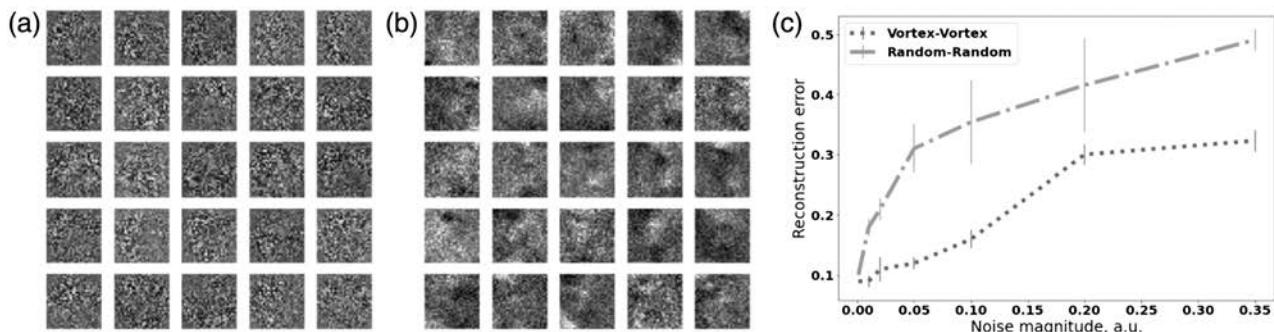
In order to estimate complexity of the pattern we employ the measure of entropy. We approximate the 2D entropy of the images using the spectra of singular value decomposition (SVD), which describes the complexity of an image. Unlike Shannon entropy [44], SVD-entropy illustrates the mixture of spatial modes that are present in an image.

We use a normalized relation for the SVD-entropy that is invariant with image intensity scaling:

$$E_{\text{SVD}} = - \sum_1^K \tilde{\sigma}_i \log_2(\tilde{\sigma}_i), \quad (10)$$

where the argument  $\tilde{\sigma}_i$  is the normalized magnitude of the singular values or the modal coefficients of the image, given as

$$\tilde{\sigma}_i = \frac{\sigma_i}{\sum_1^K \sigma_i} \quad \text{and} \quad \sum_i \tilde{\sigma}_i = 1, \quad (11)$$



**Fig. 4.** (a) Single “hot” pixel response of the random model and (b) single-pixel response of the vortex model, which demonstrates sharp edges and resolves high-contrast objects. (c) Comparison of reconstruction error for different levels of noise given high-entropy random UTS and random mask and lower SVD-entropy vortex UTS and vortex mask. This error corresponds to the scenario in which shot noise dominates the background noise.

where  $K$  is the number of singular values and  $\sigma_i$  are the singular values.

Some trends related to the SVD-entropy are illustrated in Fig. 5. If images in the set have several high singular values  $\sigma_i$ , the images may be reconstructed using fewer “elementary” patterns; those with higher entropy require many more patterns to achieve enough reconstruction accuracy. Low SVD-entropy images are smoother with fewer edges. On the other hand, images with many discontinuities exhibit a high degree of SVD-entropy.

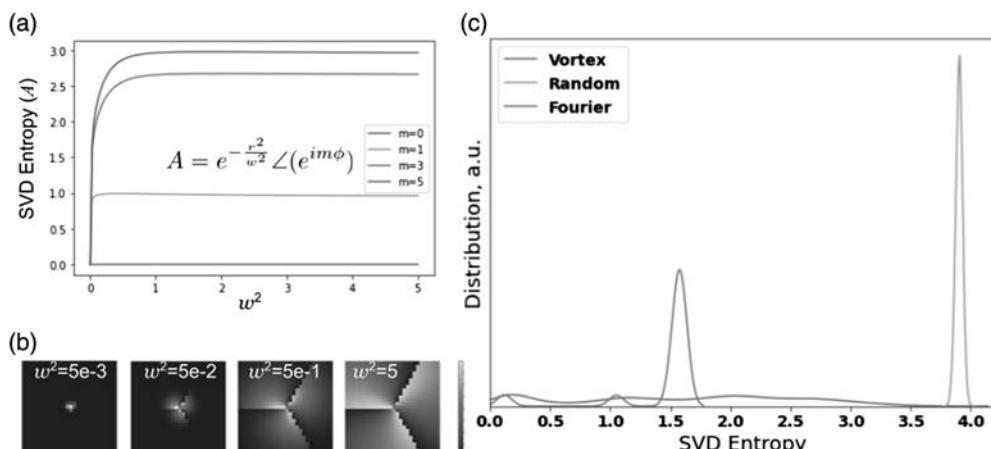
From our analysis of differently structured patterns, the SVD-entropy scales logarithmically with the edge steps or dislocations in an image [Figs. 5(a) and 5(b)]. In this illustration, we plot the phase of an  $m = 3$  vortex with varied Gaussian-beam filtering. The measure of 2D SVD-entropy aids our analysis of the UTS. The vortex UTS has a broad range and lower values of SVD-entropy in contrast to the random UTS [Fig. 5(c)].

Pertaining to our efforts toward generalized training or a UTS, we see that a low SVD-entropy training set like that with structured patterns  $X_V$  allows us to extract the structured

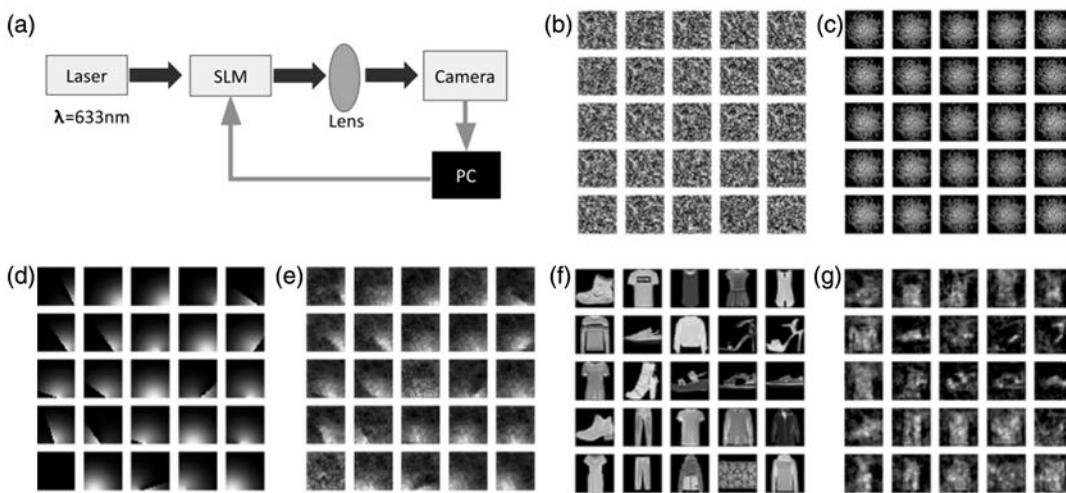
(low SVD-entropy) information from the data [Figs. 2(c), 2(f), 2(g), and 2(h)]. This effectively acts as a filter for salient features of the image. This low SVD-entropy training would be useful for some specific tasks, especially when, e.g., we are less interested in the image’s background information than in the foreground object.

## B. Heuristic Experiments

To illustrate the potential and the impact of our approach for generalizable training, we show heuristic experimental results. In simulations, almost any encoded diffraction pattern with a mask presents a learnable map for a simple neural network. However, in practice when noise is present, neural networks do not always converge. Our experimental data show that under noisy experimental conditions where light is unpolarized and the sensor data is collected with significant levels of noise, the high SVD-entropy dataset is not suitable for the task of image reconstruction: background light and distortions render a high SVD-entropy training image useless since the neural network does not learn the pattern. By contrast, a neural network trained on low SVD-entropy images is capable of recovering



**Fig. 5.** (a) SVD-entropy of a structured pattern composed of the phase of a vortex (modulus 0,  $2\pi$ ) and a Gaussian mask with radius of  $w^2$ . A few-pixel pattern has almost zero entropy, and the SVD-entropy saturates for a vortex depending on the topological charge. (b) Illustration of these patterns with  $w^2 = 5 \times 10^{-3}$ ,  $5 \times 10^{-2}$ ,  $5 \times 10^{-1}$ , and 5 corresponding to SVD-entropy values of 0.94, 1.8, 2.6, and 2.7. The SVD-entropy strongly relates to the length of the edge dislocations of an image. (c) Histogram of the SVD-entropy in the vortex  $X_V$ , Fourier  $X_F$ , and random  $X_R$  generalized training sets implemented in this project.



**Fig. 6.** (a) Schematic of experimental reconstruction with UTS. There is no spatial filter or polarizer, images are noisy, and at this wavelength, the modulation dynamic range is only  $\alpha = \pi$ . This was done intentionally to simulate poor experimental conditions with background light. (b) Sample of random UTS images and (c) sample of reconstructed images produced by random patterns, which are not learned by the simple neural network model experimentally. On the other hand, (d) simpler images with fewer edges are (e) reconstructed by the neural network. (f) Sample of ground truth images and (g) discernable reconstructed patterns when the neural network is trained by the vortex dataset.

reasonable approximations of the unseen images, as shown in Fig. 6.

Our experimental setup consists of a 633 nm helium–neon continuous-wave laser, microscope objective, HOLOEYE spatial light modulator and focusing lenses, and a CMOS 8-bit camera ( $1280 \times 1024$  pixels resolution). The setup does not include polarizers as part of the design to provide a large-background and an unmodulated signal to test the limits of image reconstruction with a simple neural network. As a result, we are unable to recover images with the zeroth-order transmitted pattern. When we instead collect the sensor data at the first diffraction maximum, we are successful with image reconstruction but only with the vortex UTS. For reconstruction purposes, small square patches of the detector pattern are taken (e.g.,  $50 \times 50$  pixels).

In our experiments with imperfect spatial beam profiles and background unmodulated noise, the simple neural networks do not converge with random masks [the results are shown in Figs. 6(b) and 6(c)]. Experimentally, we demonstrate two masks shown in Figs. 1(c) and 1(d), which are successfully learned by the neural network. The low SVD-entropy dataset composed of shapes with straight edges and curves, i.e.,  $\mathbf{X}_V$  [Eq. (8)], converges but the high SVD-entropy random  $\mathbf{X}_R$  patterns do not. Again, we find it more difficult to train a simple neural network with a high SVD-entropy UTS.

#### 4. CONCLUSION

Corners, edges, and higher-order solutions are a challenge in image reconstruction, requiring a higher degree of superposed waves [45]. This more complex representation of images is the definition of SVD-entropy in an image and suggests that the reconstruction of such images requires the learning of images composed of high SVD-entropy patterns [46]. We find, however, that this is not always the case when aiming for robust neural network-based reconstruction. In fact, generalized

training with low-entropy patterns recreates these sharp features well with edge enhancement.

We show that a simple neural network without hidden layers is capable of learning generalized image reconstruction. With this simple architecture designed to approach generalized training, it is evident that not all generalized datasets are equal. When we compare the convergence of differently structured datasets such as handwritten digits and fashion MNIST, a set of images or encoder based on vortex phase patterns (structured, low SVD-entropy, a combination of edges and curves) yields image reconstruction with lower error than a high SVD-entropy random encoder pattern that contains many edges. With a dataset such as CIFAR, the salient features are preserved in image reconstruction using a vortex UTS.

We have previously shown that a convolutional neural network can outperform a single-layer neural network but with significantly higher energy cost. The deep neural network is also less robust to noise [29]. Here, we aim to work with a “small brain” neural network rather than a deep neural network architecture. This approach has been specifically tuned with the aim of low-SWaP computational cameras. We draw the following conclusions.

- Single-layer neural networks are capable of approximating the inverse mapping from phaseless Fourier-plane intensity patterns after basic training.
- Such moderate-accuracy generalizable image reconstruction achieves high speeds (we achieve 15,000 frames per second on a 15 W laptop CPU).
- Image reconstruction with simpler neural networks is robust to the vulnerabilities and instabilities described by Ref. [28].
- Even with a simple neural network architecture and a large training basis set, we encounter differences in convergence. (Experimentally with an imperfect encoder, neural networks learn low SVD-entropy images more rapidly and reliably than high SVD-entropy.)
- Low SVD-entropy images are valuable in training neural networks to extract the salient features of the image.

Additional advantages of a UTS include what is likely a generalized upper bound for error [3], higher robustness, and high potential for low-SWaP computational cameras. Because of its low computational complexity, our approach in the future may be inverted to uncover the inverse mapping in data-driven models to solve inverse problems. A higher degree of sampling over the sensor images (i.e., zero-padding) may further reduce the reconstruction image error and even provide additional advantages, i.e., super-resolution phase retrieval from multiple phase-coded diffraction patterns [47] and depth detection [48].

**Funding.** Defense Advanced Research Projects Agency (YFA D19AP00036).

**Acknowledgment.** The authors acknowledge editing support from Ben Stewart (linkedin:benjamin-w-stewart).

**Disclosures.** The authors declare no conflicts of interest.

## REFERENCES

- Y. Park, C. Depeursinge, and G. Popescu, "Quantitative phase imaging in biomedicine," *Nat. Photonics* **12**, 578–589 (2018).
- J. Wang, J. Liang, J. Cheng, Y. Guo, and L. Zeng, "Deep learning based image reconstruction algorithm for limited-angle translational computed tomography," *PLOS ONE* **15**, e0226963 (2020).
- Y. Xue, S. Cheng, Y. Li, and L. Tian, "Reliable deep-learning-based phase imaging with uncertainty quantification," *Optica* **6**, 618–629 (2019).
- R. P. Millane, "Phase retrieval in crystallography and optics," *J. Opt. Soc. Am. A* **7**, 394–411 (1990).
- R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik* **35**, 237–246 (1972).
- J. R. Fienup, "Phase retrieval algorithms: a comparison," *Appl. Opt.* **21**, 2758–2769 (1982).
- Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE Signal Process. Mag.* **32**, 87–109 (2015).
- M. S. Asif, A. Ayremliou, A. Sankaranarayanan, A. Veeraraghavan, and R. G. Baraniuk, "Flatcam: thin, lensless cameras using coded aperture and computation," *IEEE Trans. Comput. Imaging* **3**, 384–397 (2016).
- N. Antipa, G. Kuo, R. Heckel, B. Mildenhall, E. Bostan, R. Ng, and L. Waller, "DiffuserCam: lensless single-exposure 3D imaging," *Optica* **5**, 1–9 (2017).
- K. Wakonig, A. Diaz, A. Bonnin, M. Stampanoni, A. Bergamaschi, J. Ihli, M. Guizar-Sicairos, and A. Menzel, "X-ray Fourier ptychography," *Sci. Adv.* **5**, eaav0282 (2019).
- G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nat. Photonics* **7**, 739–745 (2013).
- P. C. Konda, L. Loetgering, K. C. Zhou, S. Xu, A. R. Harvey, and R. Horstmeyer, "Fourier ptychography: current applications and future promises," *Opt. Express* **28**, 9603–9630 (2020).
- Y. Xiao, L. Zhou, and W. Chen, "Fourier spectrum retrieval in single-pixel imaging," *IEEE Photonics J.* **11**, 7800411 (2019).
- M.-J. Sun and J.-M. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: a brief review," *Sensors* **19**, 732 (2019).
- M. P. Edgar, G. M. Gibson, and M. J. Padgett, "Principles and prospects for single-pixel imaging," *Nat. Photonics* **13**, 13–20 (2018).
- M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE Signal Process. Mag.* **25**, 83–91 (2008).
- X. Hu, H. Zhang, Q. Zhao, P. Yu, Y. Li, and L. Gong, "Single-pixel phase imaging by Fourier spectrum sampling," *Appl. Phys. Lett.* **114**, 051102 (2019).
- H. Deng, X. Gao, M. Ma, P. Yao, Q. Guan, X. Zhong, and J. Zhang, "Fourier single-pixel imaging using fewer illumination patterns," *Appl. Phys. Lett.* **114**, 221906 (2019).
- Z. Liang, D. Yu, Z. Cheng, and X. Zhai, "Adaptive Fourier single-pixel imaging sampling based on frequency coefficients prediction," *Opt. Eng.* **59**, 073105 (2020).
- Z. Zhang, X. Ma, and J. Zhong, "Single-pixel imaging by means of Fourier spectrum acquisition," *Nat. Commun.* **6**, 6225 (2015).
- M. H. Seaberg, A. d'Aspremont, and J. J. Turner, "Coherent diffractive imaging using randomly coded masks," *Appl. Phys. Lett.* **107**, 231103 (2015).
- A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117–1125 (2017).
- Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–1190 (2018).
- C. A. Metzler, P. Schniter, A. Veeraraghavan, and R. G. Baraniuk, "prDeep: robust phase retrieval with a flexible deep network," in *Proceedings of the 35th International Conference on Machine Learning* (2018), pp. 3501–3510.
- E. J. Candès, X. Li, and M. Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Appl. Comput. Harmon. Anal.* **39**, 277–299 (2015).
- M. R. Kellman, E. Bostan, N. A. Repina, and L. Waller, "Physics-based learned design: optimized coded-illumination for quantitative phase imaging," *IEEE Trans. Comput. Imaging* **5**, 344–353 (2019).
- C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning requires rethinking generalization," in *Proceedings of 5th International Conference on Learning Representations* (2016), pp. 1–15.
- V. Antun, F. Renna, C. Poon, B. Adcock, and A. C. Hansen, "On instabilities of deep learning in image reconstruction and the potential costs of AI," *Proc. Natl. Acad. Sci. USA* **117**, 30088–30095 (2020).
- B. Muminov and L. T. Vuong, "Fourier optical preprocessing in lieu of deep learning," *Optica* **7**, 1079–1088 (2020).
- R. Horisaki, R. Takagi, and J. Tanida, "Learning-based imaging through scattering media," *Opt. Express* **24**, 13738–13743 (2016).
- S. Malik, A. Sardana, and Jaya, "A keyless approach to image encryption," in *International Conference on Communication Systems and Network Technologies* (2012), pp. 879–883.
- C. A. Metzler, F. Heide, P. Ranganajan, M. M. Balaji, A. Viswanath, A. Veeraraghavan, and R. G. Baraniuk, "Deep-inverse corereography: towards real-time high-resolution non-line-of-sight imaging," *Optica* **7**, 63–71 (2020).
- W. Luo, W. Alghamdi, and Y. M. Lu, "Optimal spectral initialization for signal recovery with applications to phase retrieval," *IEEE Trans. Signal Process.* **67**, 2347–2356 (2019).
- P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," *IEEE Trans. Signal Process.* **63**, 4814–4826 (2015).
- Y. Luo, D. Mengü, N. T. Yardimci, Y. Rivenson, M. Veli, M. Jarrahi, and A. Ozcan, "Design of task-specific optical systems using broadband diffractive neural networks," *Light Sci. Appl.* **8**, 112 (2019).
- E. Khoram, A. Chen, D. Liu, L. Ying, Q. Wang, M. Yuan, and Z. Yu, "Nanophotonic media for artificial neural inference," *Photon. Res.* **7**, 823–827 (2019).
- Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljačić, "Deep learning with coherent nanophotonic circuits," *Nat. Photonics* **11**, 441–446 (2017).
- G. Wetzstein, A. Ozcan, S. Gigan, S. Fan, D. Englund, M. Soljačić, C. Denz, D. A. B. Miller, and D. Psaltis, "Inference in artificial intelligence with deep optics and photonics," *Nature* **588**, 39–47 (2020).
- D. Psaltis and N. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," *Opt. Lett.* **10**, 98–100 (1985).
- S. Jutamulia and F. Yu, "Overview of hybrid optical neural networks," *Opt. Laser Technol.* **28**, 59–72 (1996).
- J. Chang, V. Sitzmann, X. Dun, W. Heidrich, and G. Wetzstein, "Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification," *Sci. Rep.* **8**, 12324 (2018).

42. J. Wu, H. Zhang, W. Zhang, G. Jin, L. Cao, and G. Barbastathis, "Single-shot lensless imaging with Fresnel zone aperture and incoherent illumination," *Light Sci. Appl.* **9**, 53 (2020).
43. S. Zheng, X. Zeng, L. Zha, H. Shangguan, S. Xu, and D. Fan, "Orthogonality of diffractive deep neural networks," arXiv:1811.03370 (2018).
44. Q. R. Razlighi and N. Kehtarnavaz, "A comparison study of image spatial entropy," *Proc. SPIE* **7257**, 72571X (2009).
45. D. Terzopoulos, "Regularization of inverse visual problems involving discontinuities," *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-8**, 413–424 (1986).
46. M. Deng, S. Li, Z. Zhang, I. Kang, N. X. Fang, and G. Barbastathis, "On the interplay between physical and content priors in deep learning for computational imaging," *Opt. Express* **28**, 24152–24170 (2020).
47. V. Katkovnik, I. Shevkunov, N. V. Petrov, and K. Egiazarian, "Computational super-resolution phase retrieval from multiple phase-coded diffraction patterns: simulation study and experiments," *Optica* **4**, 786–794 (2017).
48. Y. Hua, S. Nakamura, M. S. Asif, and A. C. Sankaranarayanan, "SweepCam—depth-aware lensless imaging using programmable masks," *IEEE Trans. Pattern Anal. Mach. Intell.* **42**, 1606–1617 (2020).

# PHOTONICS Research

## Towards smart optical focusing: deep learning-empowered dynamic wavefront shaping through nonstationary scattering media

YUNQI LUO,<sup>1,†</sup>  SUXIA YAN,<sup>1,†</sup> HUANHAO LI,<sup>2,3,†</sup>  PUXIANG LAI,<sup>2,3,4</sup>  AND YUANJIN ZHENG<sup>1,5</sup>

<sup>1</sup>School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore

<sup>2</sup>Department of Biomedical Engineering, The Hong Kong Polytechnic University, Hong Kong SAR, China

<sup>3</sup>The Hong Kong Polytechnic University Shenzhen Research Institute, Shenzhen 518034, China

<sup>4</sup>e-mail: puxiang.lai@polyu.edu.hk

<sup>5</sup>e-mail: yizheng@ntu.edu.sg

Received 20 November 2020; revised 11 January 2021; accepted 21 January 2021; posted 22 January 2021 (Doc. ID 415590); published 16 July 2021

Optical focusing through scattering media is of great significance yet challenging in lots of scenarios, including biomedical imaging, optical communication, cybersecurity, three-dimensional displays, etc. Wavefront shaping is a promising approach to solve this problem, but most implementations thus far have only dealt with static media, which, however, deviates from realistic applications. Herein, we put forward a deep learning-empowered adaptive framework, which is specifically implemented by a proposed Timely-Focusing-Optical-Transformation-Net (TFOTNet), and it effectively tackles the grand challenge of real-time light focusing and refocusing through time-variant media without complicated computation. The introduction of recursive fine-tuning allows timely focusing recovery, and the adaptive adjustment of hyperparameters of TFOTNet on the basis of medium changing speed efficiently handles the spatiotemporal non-stationarity of the medium. Simulation and experimental results demonstrate that the adaptive recursive algorithm with the proposed network significantly improves light focusing and tracking performance over traditional methods, permitting rapid recovery of an optical focus from degradation. It is believed that the proposed deep learning-empowered framework delivers a promising platform towards smart optical focusing implementations requiring dynamic wavefront control. © 2021 Chinese Laser Press

<https://doi.org/10.1364/PRJ.415590>

### 1. INTRODUCTION

Light entering a disordered medium that is thicker than a few scattering mean free paths  $l$  ( $\sim 0.1$  mm for human skin) undergoes multiple scattering due to the mismatch of the refractive index [1], leading to pervasive obstacles in communication, astronomy, and high-resolution optical delivery and imaging through or within thick scattering media, such as biological tissues. If light is coherent, scattered light along different optical paths interferes randomly, forming optical speckles, whose intensity distribution can be recorded outside the medium using cameras. Although visually random, the way that light is scattered is actually deterministic within a certain time window (usually referred to as speckle correlation time) [2]. Built upon this property, various approaches have been inspired, such as time reversal [3–6], pre-compensated wavefront shaping [2,7–13], and memory effect [1,14–16], to obtain optical focusing and imaging through scattering media. Time reversal methods, such as time-reversed ultrasonically encoded (TRUE) method [17]

and time reversal of variance encoded light (TROVE) [18], take advantage of guide stars (e.g., focused ultrasonic modulation) to encode diffused light; then, only the encoded light is time-reversed and focused inside the scattering medium. Pre-compensated wavefront shaping techniques modulate the phases of light incident into the scattering medium based on the measurement of the transmission matrix [8,10,11,19–21] or the maximization of feedback provided by the optical [7,22–25] or photoacoustic signal strength [2], with a goal to pre-compensate for the scattering-induced phase distortions. As for the memory effect, image information is encoded in the autocorrelation of the measured speckles as long as the imaging area is within the memory effect regime, and thus images can be reconstructed from speckles with iterative phase retrieval algorithms [1,26–29].

Each of the aforementioned approaches has its own advantages and limitations. For instance, pre-compensated wavefront shaping methods are attractive due to their plain working principle and experimental setup, but most reported approaches are inherently time consuming, as many iterations are required

regardless of the optimization algorithms [30,31], restricting most implementations reported thus far to static scenarios such as in fixed diffusers, which, however, scarcely exist in reality. Under the circumstance that scattering media are randomly changing or suffering from environmental disturbance that is inevitable, a focus will degrade or even vanish. To refocus light through/within time-variant media, the wavefront shaping iterations have to be repeated from the beginning each time the scattering medium changes, which is again a tedious and ineffective process [32]. This problem impedes the implementation of pre-compensated wavefront shaping from more general and realistic applications. Although imaging through non-static media has been explored with methods such as binary phase retrieval with optical phase conjugation [27,28,33,34], ghost imaging [35], shower-curtain effect [36], bispectrum analysis [37], advanced equipment [38], and memory effect [39], each has its limitations, such as the requirement for an ultrasound guide star, slow optimization, complex setup, and narrow effective regime.

Deep learning, which is a data-driven approach, has recently demonstrated wide uses to solve inverse problems like denoising [40], image reconstruction [41–46], and super-resolution imaging [47,48], owing to their superior ability in revealing complex relationships through transforming representations at one level to a higher and more abstract level [49]. The idea has also been exploited to focus light [50–52] and reconstruct images [53–55] through static scattering media. For example, Turpin *et al.* introduced neural networks for binary amplitude modulation and focused light through a single diffuser [50]; Li *et al.* trained U-Net with speckles generated by various objects with four diffusers [53]. The pre-trained network can be generalized to “unseen” objects or diffusers. All of these diffusers, however, are with the same macroscopic parameters. Sun *et al.* [56] trained five neural networks to model five different scattering conditions, and then blurred images are first classified to one of the five situations and then are fed into the pre-trained model for reconstruction. Note that, however, considering the computation time and memory budget, it is impractical to train hundreds of neural network models to cover all kinds of scattering conditions; considering only five conditions probably only get a rough classification and reconstruction.

In this paper, we aim to solve the problem comprehensively. We introduce a deep learning-empowered adaptive framework to tackle the challenge of optical focusing and refocusing through nonstationary scattering media by using wavefront shaping, which circumvents the dependency on classification or pre-trained models. A nonstationary process can be regarded as consisting of multiple piece-wise stationary stochastic processes, while the statistical properties of each stationary stochastic are analyzed to guide fine-tuning. The adaptive adjustment of hyperparameters of the proposed Timely-Focusing-Optical-Transformation-Net (TFOTNet), which is implemented by a multi-input-single-output deep convolutional long short-term memory (ConvLSTM) network, effectively circumvents the drawbacks of traditional long short-term memory (LSTM) that tends to remember only stationary variations [57]. The adaptive adjustment mechanism is non-trivial and depends on the

statistical properties of a specific stationary stochastic process, which equivalently modifies the memory units in TFOTNet. Thus, modeling the spatiotemporal non-stationarity becomes possible. Another essential of the proposed framework is the recursive fine-tuning. It makes the best leverage of the correlation between medium statuses before and after the change, which is indicated by the speckle correlation [26,58–60]. Therefore, only a small amount of newly available samples are required to fine-tune the previous network, permitting fast recovery of the focusing performance. Note that during all of the phases, the medium is generally nonstationary; it keeps changing. Although recursive fine-tuning has already allowed timely focusing recovery, adaptive recursive estimation takes it one step further, efficiently balancing the trade-off between time cost and refocusing performance, allowing controllable light delivery through the time-variant scattering medium. It is worth highlighting here that the proposed adaptive framework becomes more attractive in circumstances with fast medium motion, considerable sudden disturbance, or low signal-to-noise ratio (SNR) regarding the light refocusing performance and time consuming in fine-tuning over traditional methods.

## 2. THEORETICAL ANALYSIS OF DEEP LEARNING FRAMEWORK FOR LIGHT FOCUSING AND REFOCUSING THROUGH NONSTATIONARY SCATTERING MEDIA

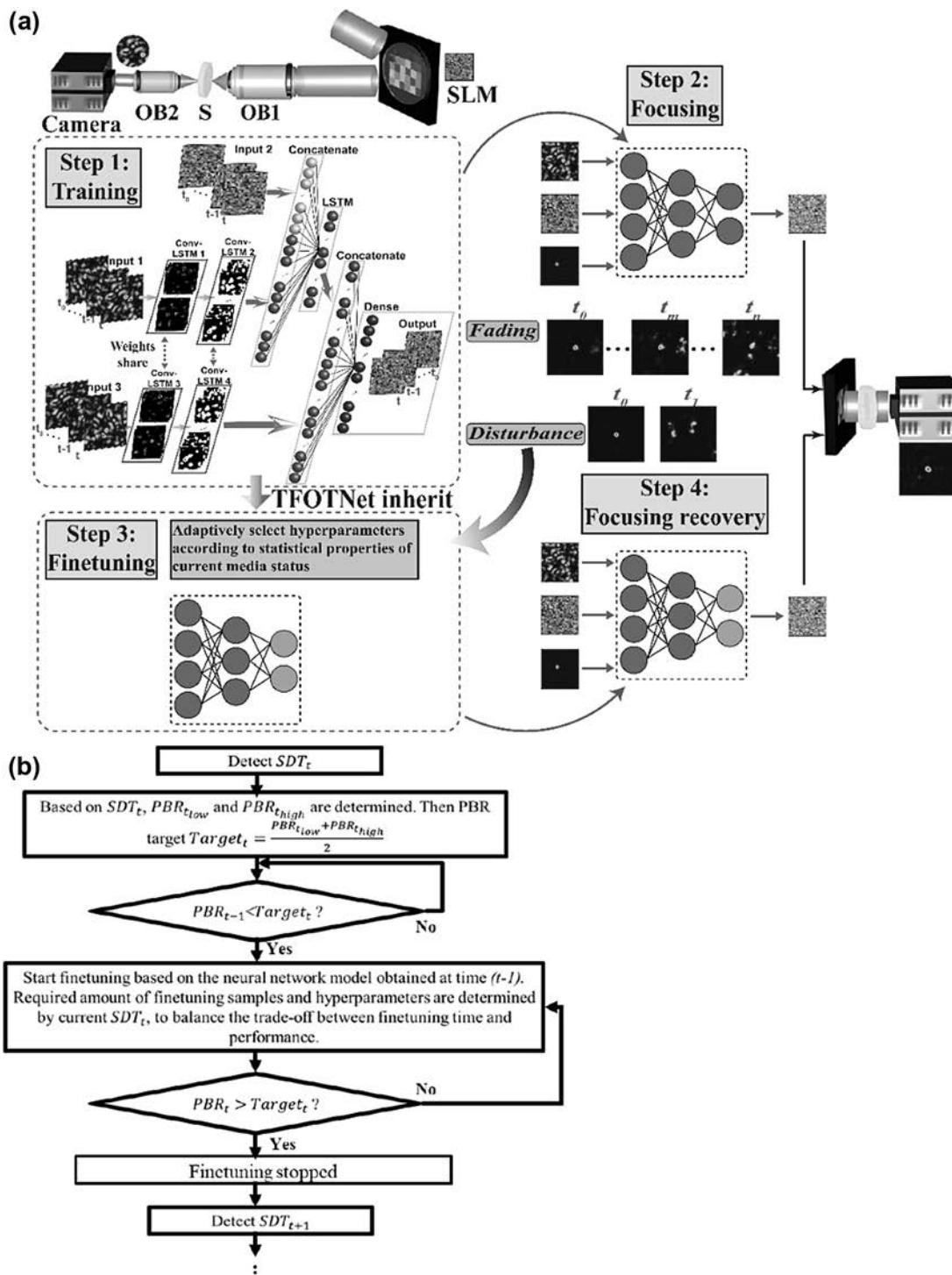
The scenario is that a monochromatic optical wave field propagates from the source to a randomly changing scattering layer at time  $t$ , and the transmitted scattered light is collected by a camera. Regular cameras only record the light intensity distribution of the speckle patterns on the receiving plane  $r_c$  (e.g., the camera plane in Fig. 1), and thus

$$I_c(t) = |E_c^{\text{out}}(t)|^2 = \left| \sum_a^N t_{ca}(t) E_a^{\text{in}}(t) \right|^2, \quad (1)$$

where  $E_a^{\text{in}}(t)$  and  $E_c^{\text{out}}(t)$  are the optical fields at  $r_a$  and  $r_c$  at time  $t$ , respectively [61], and  $r_a$  is the source plane [e.g., the spatial light modulator (SLM) plane in Fig. 1].  $t_{ca}(t)$  is a complex transmission coefficient describing light propagation from the source plane  $r_a$  to the receiving plane  $r_c$  at time  $t$ . The effects of absorption are neglected. To precisely compute the required incident complex optical field  $E^{\text{in}}(t)$  with which light is focused to the position  $r_c$  through the current scattering medium, the inverse scattering model at time  $t$  has to be obtained based on the recorded transmitted light intensity distribution  $I_c(t)$ . As seen, the inverse scattering problem is nonlinear and ill-posed, which prohibits the adoption of direct inversion methods; iterative optimization with regularization is necessary [62] to resolve this problem. The objective is to find a desired reconstruction  $E^{\text{in}}(t) = W(t)p(t)$  that minimizes the cost function formulated as

$$\arg \min_{E^{\text{in}}(t)} \|I_c(t) - |H(t)E^{\text{in}}(t)|^2\|^2 + \alpha(t) T[p(t)], \quad (2)$$

where  $H(t)$  is the forward scattering model at time  $t$ , relating the transmitted light intensity  $I_c(t)$  and the incident electrical field  $E^{\text{in}}(t)$ ,  $\alpha(t) T[p(t)]$  is the regulation,  $\alpha(t)$  is the regulation coefficient,  $W(t)$  is a convolutional transformation, and  $p(t)$  is



**Fig. 1.** Illustration of the proposed deep learning-empowered adaptive framework for wavefront shaping in nonstationary media. (a) General working principle of the proposed framework. In Step 1, samples are collected to train the TFOTNet. The structure of the proposed TFOTNet includes three inputs and one output. Input 1 is the speckle pattern, while the corresponding SLM pattern is noted as Input 2. Input 3 is the speckle pattern desired to be seen by the camera after light passes through the scattering medium in the experiment or simulation. TFOTNet output is the SLM pattern needed to get Input 3 through the present scattering medium. In Step 2, the well-trained TFOTNet can be applied to unseen speckles and output an SLM pattern that can obtain the target through the current medium. Inevitable environmental disturbance (disturbance) or non-stationary change in the medium (fading) results in degradation or even loss of the focal point. In Step 3, the pre-trained TFOTNet is fine-tuned with samples from the changing medium. Hyperparameters and fine-tuning sample amount are all adaptively chosen based on the medium status. After tuning, TFOTNet can adapt to the concurrent medium state and recover the optical focusing performance. (b) Flow chart of the proposed adaptive recursive algorithm for light focusing and refocusing in nonstationary media.

the transformation coefficients at time  $t$ .  $E^{\text{in}}(t)$  consists of  $N$  input optical modes,  $E^{\text{in}}(t) = [E_1^{\text{in}}(t) E_2^{\text{in}}(t) \cdots E_N^{\text{in}}(t)]$ .

So far, a lot of iterative algorithms have been reported to solve the inverse problems in static situations, such as the distorted Born iterative method [63], subspace optimization method (SOM) [64], and iterative shrinkage and thresholding algorithm (ISTA) [65]. Most of them rely on a building block model [66], whereas, for dynamic media, medium statuses at time  $t$  and  $t - 1$  are correlated, indicating that  $H(t)$  is not only determined by the current status but is also influenced by their previous values:

$$H(t) = g[\beta_1^{t-1} H(t-1) + \beta_2^t x(t)], \quad (3)$$

where  $\beta_1^{t-1}$  and  $\beta_2^t$  are time-dependent parameters,  $g(\cdot)$  is a non-linear function,  $H(t-1)$  is the scattering model at time  $t-1$ , and  $x(t)$  represents the information from the current scattering medium. Hence, in dynamic situations, Eq. (2) can still be solved using an iterative algorithm based on the building block model, but with temporal information included in it, and  $p(t)$  at the  $(m+1)$ th iteration is given as

$$\begin{aligned} p^{m+1}(t) &= \left\{ A_\theta \left\{ \frac{1}{L} W(t) * H(t) * I_c(t) \right. \right. \\ &\quad \left. \left. + \left[ I - \frac{1}{L} W(t) * H(t) * H(t) W(t) \right] p^m(t) \right\} \right\} \\ &= \left\{ A_\theta \left\{ \frac{1}{L} W(t) * g[\beta_1^{t-1} H(t-1) + \beta_2^t x(t)] * I_c(t) \right. \right. \\ &\quad \left. \left. + \left\{ I - \frac{1}{L} W(t) * g[\beta_1^{t-1} H(t-1) + \beta_2^t x(t)] \right. \right. \right. \\ &\quad \left. \left. \left. * g[\beta_1^{t-1} H(t-1) + \beta_2^t x(t)] W(t) \right\} p^m(t) \right\} \right\}, \end{aligned} \quad (4)$$

where  $L(t)$  is the Lipschitz constant at time  $t$ ,  $L(t) \leq \text{eig}[W(t) * H(t) * H(t) W(t)]$ . From Eq. (4), the iterative optimization process can be regarded as a sequence of linear filtering by kernel  $I - [1/L(t)]W(t) * H(t) * H(t)W(t)$  and bias  $[1/L(t)]W(t) * H(t) * I_c(t)$ , followed by a point-wise nonlinear operation  $A_\theta$  by value  $\theta$ , and  $I$  is the identity matrix. Meanwhile, information from previous medium statuses is transmitted over time and imposes influence on the inverse scattering model at time  $t$ . Equations (2)–(4) suggest that TFOTNet, which is a multi-input-single-output ConvLSTM network, is suitable for solving the inverse scattering problems in dynamic situations. The inverse scattering problem is formulated as a regression task, while the learning process is evaluated by the mean integrated squared error [67,68]:

$$\text{MISE} = E \left\{ \int_0^T \iint_{S_a} [y_t(r_a) - \hat{y}_t(r_a)]^2 d^2 r_a dt \right\}, \quad (5)$$

where  $y_t(r_a)$  and  $\hat{y}_t(r_a)$ , respectively, denote the model prediction and true values at time  $t$ ,  $r_a$  is a segment on the SLM plane  $S_a$ , and  $E$  denotes the expected value. To resolve wavefront shaping problems,  $y_t(r_a)$  and  $\hat{y}_t(r_a)$  are the predicted and true phase values of the incident optical mode  $E_a^{\text{in}}(t)$ , respectively.

The speckle correlation theory in random media suggests that, when the configurations of the scatterers are changed

randomly, the scattering media before and after moderate change are correlated [69]. For dynamic media whose properties are time-variant, both spatial and temporal speckle correlations exist, and thus the speckle correlation is shown as [70]

$$\begin{aligned} C(t-t', r-r') &= \frac{\langle I(t, r)I(t', r') \rangle - \langle I(t, r) \rangle \langle I(t', r') \rangle}{\langle I(t, r) \rangle \langle I(t', r') \rangle} \\ &= C_1(t-t', r-r') + C_2(t-t', r-r') \\ &\quad + C_3(t-t', r-r'). \end{aligned} \quad (6)$$

According to the results reported by Feng *et al.* [69], the intensity correlation function  $C(t-t', r-r')$  can be regarded as consisting of three contributions  $C_1(t-t', r-r')$ ,  $C_2(t-t', r-r')$ , and  $C_3(t-t', r-r')$ , governing the short-range correlation, long-range correlation, and infinite-range correlation, respectively [71]. For most scattering media, the magnitude of  $C_1(t-t', r-r')$ ,  $C_2(t-t', r-r')$ , and  $C_3(t-t', r-r')$  decreases in sequence, but also decays more slowly with the increase of gap between  $t$  and  $t'$  or  $r$  and  $r'$  [70]. The proposed framework encodes the correlation between medium statuses, propagating the information over time; as a consequence, an accurate inverse model can be constructed.

### 3. RESULTS

#### A. Working Principle

The structure of the proposed TFOTNet is shown in Fig. 1(a). TFOTNet has three inputs and one output. Inputs 1 and 2 are paired, while Input 3 and the output are paired. Referring to Fig. 1(a), light is firstly reflected by the SLM, and its phase pattern is adjusted by the SLM; thus the optical phase patterns are represented by the SLM patterns. After the SLM, light will go through the diffuser and be scattered, forming speckles. The intensity distribution of the training speckle patterns will be recorded outside the diffuser by the camera, which is Input 1. The corresponding SLM pattern is Input 2. This forms a mapping from the training speckle pattern (Inputs 1) to the trained SLM pattern (Inputs 2), and it acts as a regularization term. Incorporating this regularization input into the TFOTNet, the targeted relationship from Input 3 to the output is obtained, and it is used to resolve the inverse scattering problems in real time based on the regularized cost function Eq. (2). Input 3 is the speckle pattern desired to be seen by the camera after light passes through the scattering medium in the experiment or simulation. The output of TFOTNet will be the corresponding SLM pattern that can lead to Input 3.

Inverse scattering problems are ill-posed, which may lead to difficulties in neural network training [72]. Offering prior information to regularize the inverse problem can mitigate the burden in training, which plays a significant role in successfully resolving inverse problems [73,74]. Besides setting analytic priors manually, it has also been reported that prior terms can be directly learned during the training of neural networks, which is tailored to the statistics of the training images, indicating a stronger regularization [75,76]. Chang *et al.* adopted an adversarial method to jointly train two networks, where one offers prior information, while the other one conducts inverse projection [76]. Inspired by these, the proposed TFOTNet consists of two parts: prior knowledge about scattering provided

by Inputs 1 and 2, and inverse mapping from Input 3 to the output. Through training, the network learns to extract suitable priors from Inputs 1 and 2, and they are passed to facilitate the resolving of the inverse problem represented by Input 3 and the output, alleviating the training burden and improving the modeling accuracy when compared with methods that directly learn the inverse mapping without any other knowledge.

Generally, in transfer learning, only the last few layers, rather than whole neural networks, are fine-tuned [77,78], as the last layers are task specific, while the earlier ones are modality specific [79]. Information learned by earlier layers can be shared among all inverse scattering problems, while the last few layers are customized for adapting to specialized changing conditions. Therefore, when the TFOTNet needs to be fine-tuned in the experiment, only the last layer in the TFOTNet is adjusted, while all other layers are frozen. By doing so, both time and computational resources can be saved without significant sacrifice of accuracy. The two ConvLSTM layers, ConvLSTM1 and ConvLSTM2, extract and abstract image features from Input 1; meanwhile, they pass the useful features from previous statuses throughout the network. Then, these features are flattened to concatenate with Input 2, which has also been flattened. The combination serves as the input to the first LSTM layer, followed by a dropout layer. The outputs of the LSTM layer concatenate with the features gathered from Input 3. The final TimeDistributed dense layer predicts the SLM pattern needed for Input 3 in the current situation. ConvLSTM1 and ConvLSTM2 consist of 16 and 32 filters, respectively, and the filter size of each layer is  $7 \times 7$  and  $5 \times 5$  with the stride setting as  $3 \times 3$  and  $2 \times 2$ , respectively. ConvLSTM1 and ConvLSTM3, ConvLSTM2 and ConvLSTM4 share the same structure and weights, respectively. The number of neurons in the LSTM layer is 256 with a dropout rate set to 0.3. The number of neurons in the output layer is the same as the size of the SLM patterns. Kernel initializers of all layers are set as glorot normal. Mean squared error is employed as the loss function. Adam is used as the optimizer with alpha, beta1, beta2, and epsilon set as 0.0005, 0.9, 0.99, and 0.0001, respectively. It is worth noting that the proposed TFOTNet is a general network that can be applied to deal with speckle images and SLM patterns of arbitrary size. Herein, we just introduce a specific implementation for our typical setup as a proof-of-concept. The output size of TFOTNet is determined by the size of the SLM patterns, which is user defined. Kernel size is adjusted in the light of the relative size of speckle grains and recorded speckle images. In general, with smaller speckle grains, both the kernel size and stride have to be reduced accordingly. For larger speckle images or SLM patterns, naturally, more training and fine-tuning samples are required; meanwhile, the number of neurons will be increased, and the dropout rate is enlarged as well to avoid overfitting. The activation function of all layers is tanh, except for the last output layer whose activation function is sigmoid. The recurrent activation function of all ConvLSTM layers is set as hard sigmoid. The TensorFlow Keras library is used to construct the model.

It is worth noting here that the lynchpin of the proposed framework is the adaptive recursive fine-tuning system, rather

than any other specific implementations. Despite this, TFOTNet includes information from Inputs 1 and 2 to facilitate the ill-posed inverse mapping from Input 3 to the output; thus, it not only allows for more efficient modeling over conventional single-input-single-output ConvLSTM network or convolutional neural network (CNN) (both simulation and experimental comparisons are shown in the following content), but is also a general network whose structure is scalable to accommodate various applications. Considering that SLMs are widely used to modulate incident optical wavefronts, as shown in Fig. 1(a), in this article, we employ SLM patterns to represent the incident optical phase patterns.

The working flow of the proposed adaptive deep learning framework for light focusing and refocusing in nonstationary media is illustrated in Fig. 1(a). First, samples are collected for TFOTNet training and initialization. After that, the well-trained TFOTNet is able to establish an inverse scattering model statistically that can accurately map the intensity distribution of speckles to their corresponding SLM patterns. Then, the desired speckle (a preset focused speckle pattern is used here) is sent to the TFOTNet through Input 3, and the TFOTNet outputs the SLM pattern that is required to restore the desired pattern for the current scattering system. Considering that the scattering media are nonstationary, and environmental perturbations with time are inevitable, an optical focus can be faded or even lost. To cope with it, ad hoc samples from the real-time medium are offered to recursively fine-tune the TFOTNet that is obtained previously. Meanwhile, hyperparameters are all adaptively chosen according to the instant status of the medium. During the fine-tuning phase, only the weights of the last layer in the TFOTNet are adjusted while all other layers are frozen. After the directed adjustment, the fine-tuned TFOTNet will be able to produce an SLM pattern to recover the focusing performance in a short period of time.

Figure 1(b) elaborates the proposed adaptive recursive algorithm to handle the spatiotemporal non-stationarity. In the whole article, light focusing performance is quantitatively evaluated by the peak-to-background-ratio (PBR), which is defined as the ratio between the intensity of the focal point and the mean intensity of the background [80]. Medium changing speed is characterized by speckle decorrelation time (SDT), which is defined as the time duration that the intensity auto-correlation function decreases from 1 to  $1/e$  of its initial value [81]. Smaller SDT corresponds to faster medium altering speed. At time  $t$ , the SDT of the current medium state is computed, and the PBR target,  $\text{Target}_t$ , is determined based on the SDT (the method to calculate the PBR target is elaborated in Section 4). The PBR target indicates the pre-defined PBR that targets to be attained after light refocusing. The adaptive PBR target is employed to balance the trade-off between the fine-tuning cost and the focusing recovery performance. More samples are needed to enhance the PBR to a pre-defined level when the scattering medium changes faster, suggesting that longer time is required; thus focusing tracking performance will be affected. With an adaptive PBR target, a faster SDT accommodates to a relatively lower PBR target, which needs fewer fine-tuning samples, shortening the fine-tuning time. Instant PBR is

compared with Target<sub>t</sub>, and fine-tuning will not be initialized until the instant PBR is lower than Target<sub>t</sub>. Pairs of the SLM pattern and the corresponding speckle pattern are collected during the changing process of the scattering medium for fine-tuning, and the required fine-tuning sample amount and hyperparameters of the network are all chosen based on the SDT. The influence of hyperparameters on fine-tuning time is discussed in Section 4. The recursive algorithm indicates that the fine-tuning is based on the network obtained at time  $t - 1$ , which can make the best use of the speckle correlation; thus, time cost in fine-tuning can be significantly reduced when compared with traditional iterative algorithms. Once the instant PBR after fine-tuning is higher than Target<sub>t</sub>, fine-tuning will be ceased. With iterations of such an adaptive recursive fine-tuning process, optical focus can be recovered from deterioration in time, allowing for maintenance of a focal point with acceptable performance. Proof-of-concept simulation and experimental results with a typical setup are shown below as a verification. It should be highlighted here that all results demonstrated in the article, such as the PBR target and the amount of fine-tuning samples, are valid in all conditions with proper scaling in the light of specific implementations, not just limited to the setup used here. The methods of scaling will be discussed in the experiments part.

## B. Simulation and Experimental Results

### 1. Simulation Results

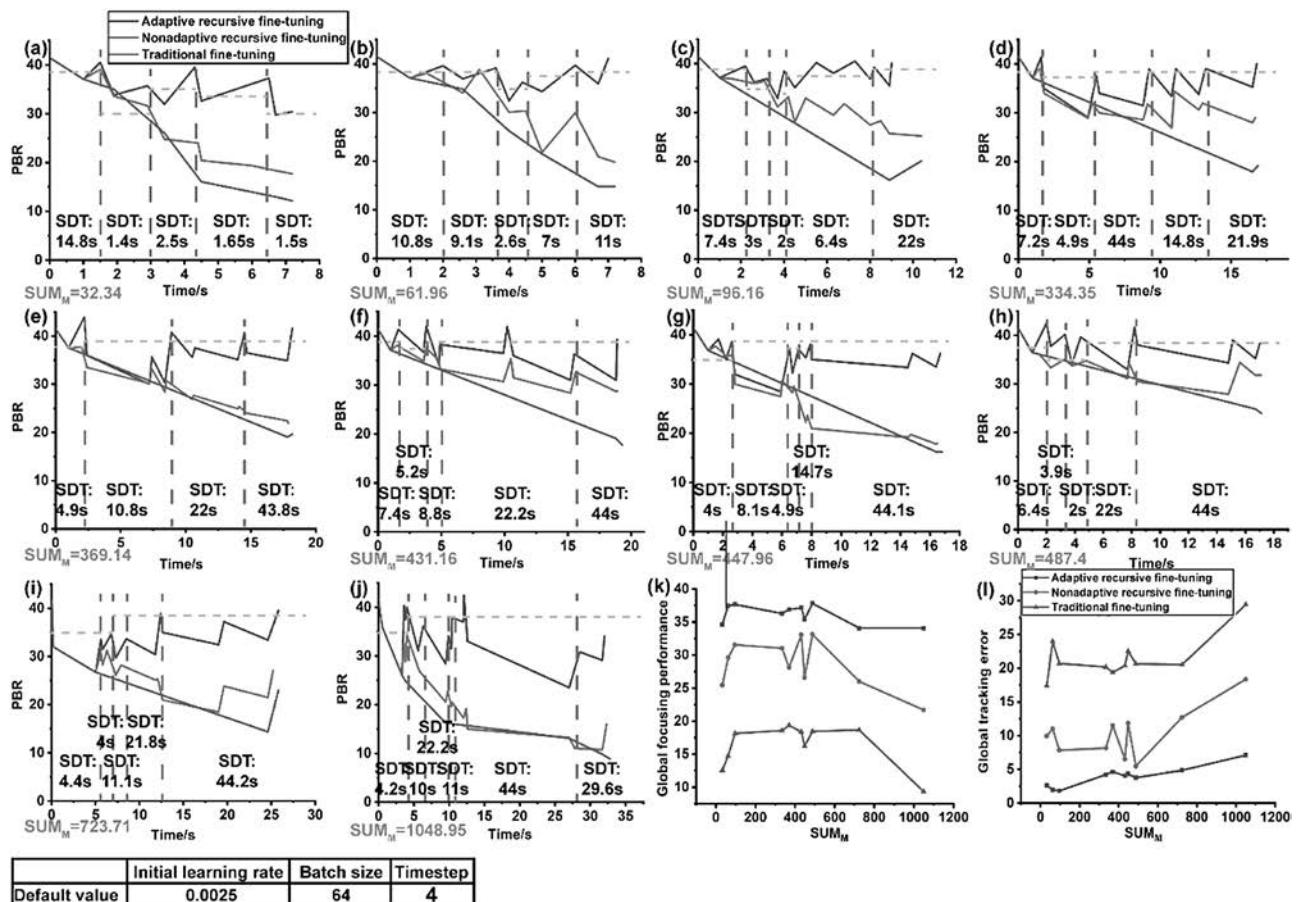
Proof-of-concept continuous nonstationary processes are simulated to clearly manifest the effect of the adaptive recursive fine-tuning algorithm in dealing with spatiotemporal nonstationarity. One nonstationary course can be regarded as consisting of multiple piece-wise stationary stochastic sub-processes characterized by different SDT. Meanwhile, the time duration of each stationary sub-process also varies. To simulate the scattering process, a transmission matrix TX( $t$ ) is used to describe a disordered medium at time  $t$ , following a circularly symmetric Gaussian distribution [23]. For a medium that is not static, the medium status at time  $t + \Delta t$  is represented by TX( $t$ ) +  $\Delta$ TX( $\Delta t$ ), and  $\Delta$ TX( $\Delta t$ ) also follows a circularly symmetric Gaussian distribution.  $\Delta$ TX( $\Delta t$ ) of different variances is employed to model media of various altering speeds. The size of the SLM patterns is set as  $32 \times 32$ , while the size of the speckle patterns is  $64 \times 64$ . First, in Step 1, a total of 10,000 samples are created for TFOTNet initialization and training. Sample collection time is estimated using the maximal frame rate of a commercial liquid crystal on silicon (LCoS) SLM, which is generally 60 Hz, and the SDT of the medium is as long as 10 min, resulting in correlation between medium statuses when the training sample collection initiates and ends reaching 0.8. After training, a desired speckle pattern [as shown in Fig. 1(a)] is sent to the well-trained TFOTNet, and a focused speckle can be obtained with the predicted SLM pattern. The PBR of the focused speckle obtained with the original trained model is 41.5. Since the proposed network is scalable, with the

SLM pattern or speckle image size becoming larger, it can be expected that PBR will also increase. It is worth noting that the sample collection speed can be expedited nearly 400 times if faster modulators such as a digital micromirror device (DMD) are applied to conduct wavefront modulation whose frame rate can reach 23 kHz [82].

However, the PBR is decreasing, and the focal point is fading over time, as the scattering medium is altering over time, calling for the necessity of fine-tuning after a certain period of time. Herein, 10 nonstationary processes are randomly generated. Three fine-tuning algorithms, i.e., the adaptive recursive algorithm, nonadaptive recursive algorithm, and traditional algorithm, demonstrate significant differences in the PBR recovering performance, which is discussed separately in the following. Simulated focusing recovery results with three algorithms using TFOTNet in 10 nonstationary processes are shown in Figs. 2(a)–2(j), where the same legends are used. For each steadily altering sub-process, its PBR target (the ideal case) is indicated by the yellow dashed line. Six SDT intervals are used, and their corresponding PBR target is given in Table 1, serving as a criterion to evaluate the focusing recovery performance (details about the calculation of an adaptive PBR target are elucidated in Section 4). Although different setups result in different initial focused speckles, the ratio between the PBR target and PBR of the initial focused pattern remains unchanged as long as the SDT is the same. By doing so, the presented results can be safely scaled to any other implementation. With the adaptive recursive algorithm, the hyperparameters are selected on the basis of statistical properties of each stationary variation; with the nonadaptive recursive algorithm, the hyperparameters remain as the default values (shown in Fig. 2) all the time, regardless of the SDT changes. For a fair comparison, each time adaptive fine-tuning is conducted the same samples are offered to the nonadaptive recursive algorithm to do one fine-tuning as well. As for the traditional algorithm, it is not a recursive one; instead, traditional wisdoms only conduct fine-tuning once in the last sub-process from the original trained model due to the lack of engines of sensing intermediate processes, the hyperparameters remain as the default values, and total sample amount is the same as those in the recursive methods. It can be seen from Figs. 2(a)–2(j) that, qualitatively, PBR achieved by the adaptive recursive approach (gray line) is always the highest, while the traditional fine-tuning algorithm (blue line) demonstrates the worst performance. In view of the structure of TFOTNet, which consists of multiple ConvLSTM cells, adjusting those hyperparameters equivalently modifies the memory units in effect. By analyzing the statistical properties (including mean value and autocovariance) of diverse stationary variations, hyperparameters can be adjusted adaptively, compensating for the limitations of the conventional ConvLSTM network that it lacks the capability to encode high-order nonstationary spatiotemporal variations [57].

**Table 1. SDT-Dependent PBR Target**

SDT/s	SDT $\geq$ 7.4	7.4 > SDT $\geq$ 4.4	4.4 > SDT $\geq$ 2	2 > SDT $\geq$ 1.6	1.6 > SDT $\geq$ 1.4	1.4 > SDT
PBR target	38.75	37.5	35	33.75	30	27.5



**Fig. 2.** Fine-tuning results with ten random nonstationary processes using three different algorithms. The 10 nonstationary processes can be regarded as consisting of multiple piece-wise stochastically stationary sub-processes, while the SDT and time duration of each sub-process are different. (a)–(j) Fine-tuning results with the adaptive recursive algorithm (gray line), nonadaptive recursive algorithm (red line), and traditional fine-tuning algorithm (blue line) in the 10 nonstationary processes. Each process is characterized by  $SUM_M$ , which is the sum of the product of SDT of each sub-process and its time duration. Figures (a)–(j) use the same legend. (k) Global focusing performance of three fine-tuning algorithms in the ten random nonstationary processes. (l) Global tracking error of three fine-tuning algorithms in the ten random nonstationary processes. Figures (k) and (l) use the same legend. The inserted table lists the default values of these hyperparameters used in simulation. Specific hyperparameters and fine-tuning sample amounts used in the simulation are available in Ref. [83].

Quantitatively, one nonstationary process is characterized by the sum of the product of SDT of each sub-process and its time duration:

$$SUM_M = \sum_{j=1}^M SDT(j) \times Duration(j), \quad (7)$$

where  $M$  represents the total number of stationary stochastic sub-processes contained in a nonstationary course. Lower  $SUM_M$  values suggest smaller SDT, shorter time duration, or both. To quantitatively evaluate the overall light focusing maintenance performance, the mean value of PBR over the whole nonstationary process is adopted:

$$GFP \text{ (global focusing performance)} = \sum_{i=1}^N PBR(i)/N. \quad (8)$$

With the three algorithms, PBRs at the time when each fine-tuning starts and ends are recorded, and then all of the recorded PBR values are averaged, serving as the global focusing

performance (GFP). Therefore,  $N$  is not a fixed value; instead, it is determined by how many times fine-tuning has been conducted in a nonstationary course. It is obvious that the higher the GFP, the better the focusing across time. The GFP of the three algorithms in these 10 nonstationary processes is illustrated in Fig. 2(k). Figures 2(k) and 2(l) use the same legend. In all cases, the adaptive recursive algorithm demonstrates the highest GFP, ranging from 34 to 38, while the traditional fine-tuning algorithm shows the worst results with GFP fluctuating between 9 and 19. The GFP enhancement percentage achieved by the adaptive and nonadaptive recursive algorithms over the traditional one is 82%–264% and 39%–132%, respectively, which is calculated by GFP enhancement percentage =  $\left( \frac{\text{GFP with adaptive or nonadaptive recursive algorithm}}{\text{GFP with traditional algorithm}} - 1 \right) \times 100\%$ . These results manifest the merits of the adoption of the recursive approach, where they can make the best use of both temporal and spatial correlations of the medium for fine-tuning, allowing much better focusing recovery. Meanwhile, the adaptive

algorithm realizes 12%–57% GFP enhancement over the non-adaptive recursive algorithm, confirming that through the adaptive adjustments of hyperparameters, spatiotemporal non-stationarity can be better learned by neural networks. As for the traditional algorithm, medium status at the time when fine-tuning is conducted has little correlation with the status at the time when the initial model is obtained ( $t = 0$ ), which will definitely lead to poor fine-tuning performance.

In addition, the root mean squared error with respect to the adaptive PBR target is employed to measure the tracking performance of an algorithm over the whole nonstationary processes:

GTE (global tracking error)

$$= \sqrt{\frac{\sum_{i=1}^N [\text{PBR}(i) - \text{PBR target}(i)]^2}{N}} \quad (9)$$

Small global tracking error (GTE) means that over the whole process, the PBR fluctuates mildly around the target, indicating an accurate and timely tracking of the optical focusing. The GTE of the three algorithms in these 10 non-stationary processes is illustrated in Fig. 2(l). As seen, the adaptive recursive algorithm shows the lowest values (1.8–7.1), while the traditional algorithm shows the largest values (17–29), as fine-tuning is conducted without considering process variations. The reduction percentage in GTE achieved by the two recursive algorithms over the traditional one is computed by GTE reduction percentage =  $\left(1 - \frac{\text{GTE with adaptive or nonadaptive recursive algorithm}}{\text{GTE with traditional algorithm}}\right) \times 100\%$ . As seen, the adaptive and nonadaptive recursive methods reduce the GTE over the traditional algorithm by 76%–91% and 37%–73%, respectively, suggesting that recursive algorithms demonstrate much better focusing tracking performance. Meanwhile, the adaptive recursive algorithm realizes 31%–81% reduction in GTE over nonadaptive recursive results. Due to the influence from environmental disturbance, PBR may drop sharply [as seen in Figs. 2(d)–2(j)], indicating that more time is required to recover the PBR. Therefore, the increase of GTE in these circumstances is inevitable, as shown in Fig. 2(l). In situations without perturbations [as shown in Figs. 2(a)–2(c)], PBR can be well maintained to be around the target with the proposed adaptive recursive algorithm, and the GTE is as low as 1.8–2.6. It should be noted that the conclusions deduced above are effective for all implementations of the proposed deep learning-empowered adaptive framework, not just limited to the setup and random nonstationary processes demonstrated in the article. For all realizations, similar performance improvements can be obtained.

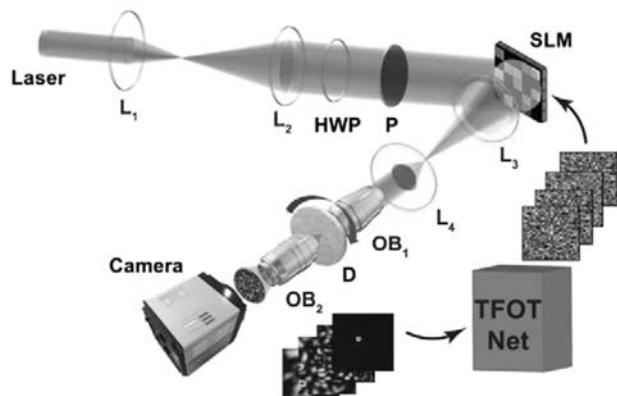
It is worth noting that the fine-tuning time can be significantly reduced, and the SDT can be much smaller than values shown here if faster modulators and/or more powerful computation engines are adopted. As a proof-of-concept, in the simulation, the sample collection speed is estimated based on the maximal frame rate of commercially available LCoS-SLM, which is 60 Hz. As for computation, the TensorFlow Keras library is adopted, and the computing unit is an Acer Predator G9-792, 16 Gb RAM, and a GTX 980M graphics

processing unit (GPU). However, if a DMD is applied to conduct wavefront modulation, whose frame rate can reach 23 kHz [82], together with onboard data acquisition, sample collection can be expedited by nearly 400 times. Furthermore, if a more powerful GPU or workstation such as the Nvidia Tesla series is employed, the computation speed will be improved by at least three times. Thus, the fine-tuning process can be speeded up by nearly 1000 times, indicating that the proposed framework becomes likely to achieve wavefront shaping in dynamic situations, such as *in vivo* tissues that decorrelate as fast as several milliseconds [84].

In Fig. 2(i), due to the sharp PBR drop at the beginning (from 41.5 to 26.7) and fast medium change (SDT = 4.4 s), even though recursive fine-tuning is conducted, the target PBR cannot be reached. Nevertheless, one attractive property of the adaptive method is that once a slower changing sub-process is detected, it is capable of making up the earlier PBR loss. As seen, with the fourth sub-process whose SDT has increased to 21.8 s, PBR is enhanced to meet the target, reaching 38.9. In contrast, the refocusing ability of nonadaptive algorithms is exacerbated, and they never meet the PBR target. Nonadaptive algorithms lack the ability to sense the current situation and make adjustments accordingly. Instead, it applies the same fine-tuning system to all processes regardless of their SDTs, which will definitely result in modeling deficiency.

## 2. Experimental Results

After the verification with simulations, experiments are conducted. The experimental setup is illustrated in Fig. 3. Light emitting from a He–Ne CW laser (633 nm, Melles Griot) is expanded by a telescope by 4.3 times. Then, a half-wave plate and a polarizer are followed to adjust the polarization of the incident light to be parallel to the long axis of an SLM (X13138-01, Hamamatsu). The light wavefront is modulated by the SLM, after which light passes through two successive lenses and is focused onto the surface of a diffuser (ground glass of 120 Grit, Edmund) by an objective lens (TU Plan Fluor

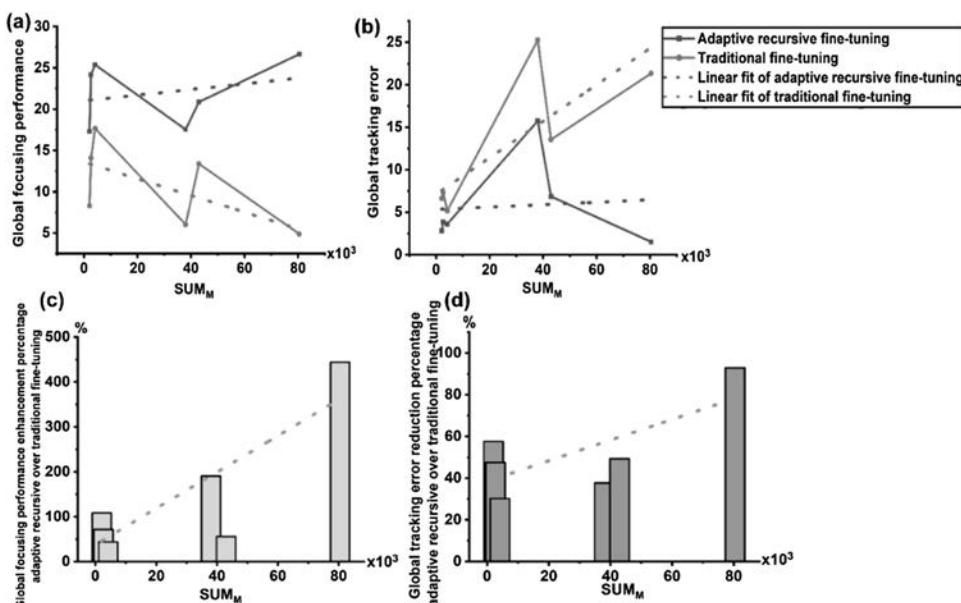


**Fig. 3.** Schematic of the experimental setup. Light is expanded by two lenses ( $L_1$  and  $L_2$ ), and then a half-waveplate (HW) and a polarizer (P) adjust the polarization state of the light incident onto the spatial light modulator (SLM). Light is modulated and reflected by the SLM, then passes through two lenses ( $L_3$  and  $L_4$ ), and is focused onto a diffuser (D) surface by an objective lens ( $OB_1$ ). Scattered light is collected by another objective lens ( $OB_2$ ) and recorded by a camera.

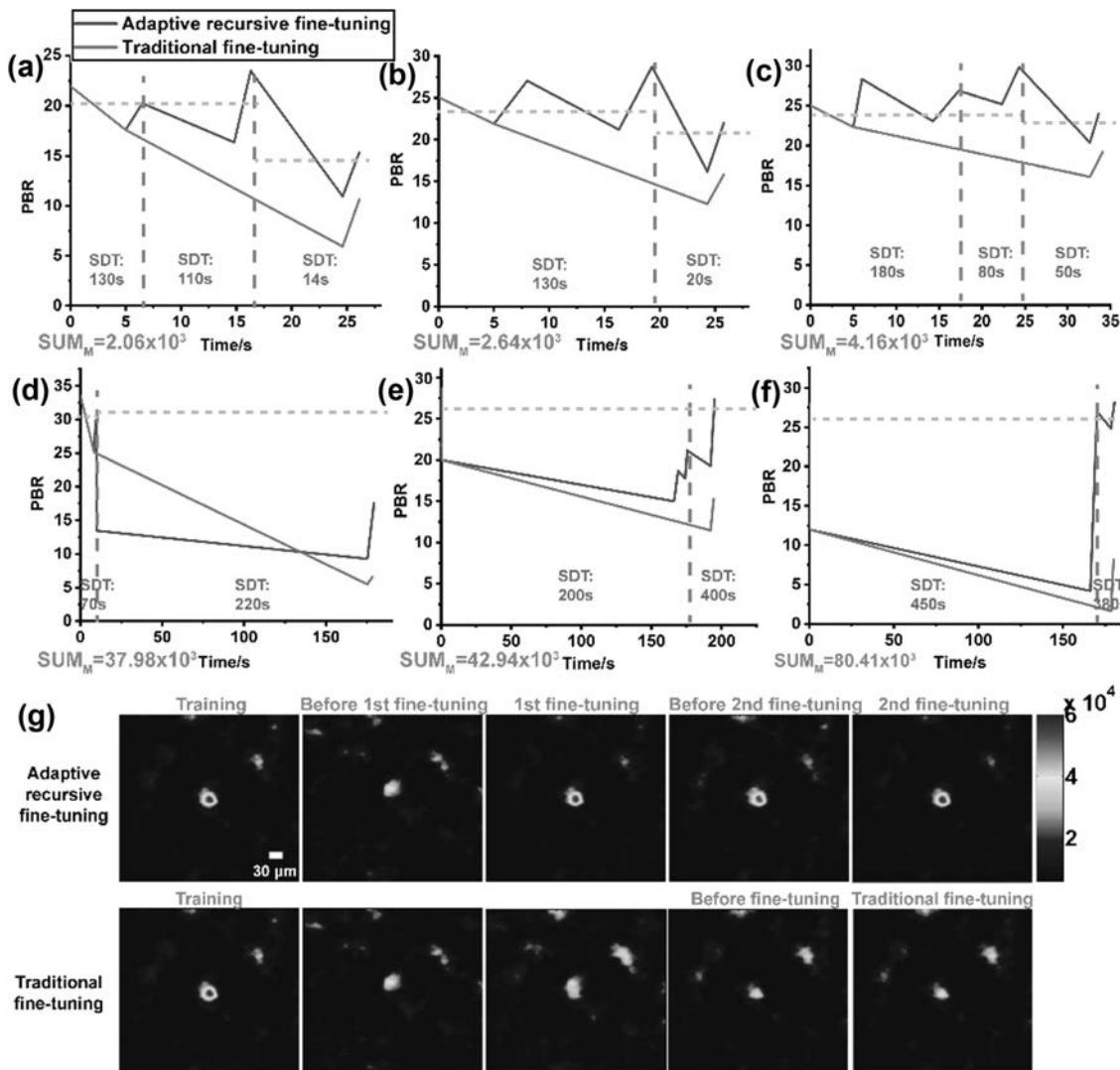
50 $\times$ /0.80, Nikon). The light undergoes multiple scattering from the diffuser, and then the scattered light is collected by another objective lens (TU Plan Fluor 20 $\times$ /0.45, Nikon) placed behind the diffuser. Finally, the speckles are recorded via a camera (Zyla s4.2, Andor). The resolution of the SLM screen is 1280  $\times$  1024, and it is divided to 32  $\times$  32 macropixels to display the SLM patterns, i.e., one macropixel contains 40  $\times$  32 pixels. The dimensions of the speckle patterns recorded by the camera are 64  $\times$  64 pixels. In experiment, we use 32 gray steps in the SLM to represent phase values from 0 to  $2\pi$ . Due to the limitation of precision from the rotating stage (Motorized Precision Rotation Stage PRM1/MZ8, Thorlabs), the diffuser is rotated once as 100 samples are collected, and the equivalent rotating speed varies from 2.5 to 10 mdeg/s. Note that the nominal frame rate of the SLM is 60 Hz. However, due to the existence of the rising/falling transition time of the SLM, as well as limitations posed by the camera exposure time and the transmission speed between the laptop and the system, the frame rate achieved in operation is only  $\sim$ 6 Hz. This has restricted the medium change speed demonstrated in the current phase of the experiment. The diffuser is rotating at different speeds to create various SDTs. As the frame rate is 10 times slower than that used in simulation estimation, SDT in simulation has to be enlarged 10 times accordingly to be consistent with the experimental conditions. For instance, in the experiment, the ratio between the PBR target and the initial PBR in a stationary stochastic process whose SDT is 200 s should be the same as that in the stationary variation whose SDT is 20 s in simulation. However, it is worth noting that there is no fundamental limitation on the speed and performance of the proposed framework if faster modulators such as DMD can be applied, and hence the results demonstrated here are scalable.

Three experiments with and without environmental disturbance are conducted, respectively, and, in each experiment, the adaptive recursive algorithm and traditional algorithm are investigated for comparisons. The proposed framework consists of two parts: recursive fine-tuning and adaptive adjustments of hyperparameters. It has been proved by simulation that hybridizing these two systems will demonstrate better performance in handling non-stationarity rather than only employing recursive fine-tuning; thus the nonadaptive recursive algorithm is not applied to experiments. For a fair comparison, the total fine-tuning samples of the adaptive recursive algorithm and traditional algorithm during a nonstationary process are the same, and both algorithms are implemented with TFOTNet. Experimental results are shown in Figs. 4 and 5. Results without environmental perturbations are shown in Figs. 5(a)–5(c), while results with disturbance are given in Figs. 5(d)–5(f). The same legend is used in these figures. Figures 4(a) and 4(b) indicate the GFP and the GTE of the six experiments, respectively. The SDT of each sub-process is shown in the figures, while the PBR target (the ideal case) is indicated by the yellow dashed line. In all experiments, the first step is TFOTNet initialization and training using 10,000 samples to obtain a focused speckle, and the initial PBR is displayed in the figures at  $t = 0$ . As stated, the PBR target is determined by the SDT as well as the PBR of the initial focused speckle. For media of the same SDT, the ratio between the PBR target and the initial PBR always remains the same. Thus, the PBR target can be deduced from the typical simulation results demonstrated above.

As seen from Figs. 5(a)–5(c), under circumstances without sudden disturbance, the PBR target can always be reached after fine-tuning using the adaptive recursive algorithm (gray line), while the traditional algorithm (red line) never reaches the



**Fig. 4.** Experimental results. (a) Global focusing performance in the six experiments with the adaptive recursive algorithm and traditional algorithm. (b) Global tracking error in the six experiments with adaptive recursive algorithm and traditional algorithm. Figures (a) and (b) use the same legend. (c) The enhancement percentage in global focusing performance achieved by the adaptive recursive algorithm over the traditional algorithm. (d) The reduction percentage in global tracking error achieved by the adaptive recursive algorithm over the traditional algorithm.



**Fig. 5.** (a)–(c) Experimental results of the three trials without environmental disturbance. The SDT of each stationary sub-process is shown in the figure, and the PBR target (ideal case) is indicated by yellow dashed lines. (d)–(f) Results of the three experiments with environmental disturbance. Figures (a)–(f) use the same legend. (g) Speckle images recorded during a nonstationary process with environmental perturbation using adaptive recursive and traditional algorithms. In (g), all speckle images use the same colormap and scale and are interpolated to  $253 \times 253$  for a better view. The color bars indicate the detected light intensity in arbitrary units. The middle image in the bottom row (traditional fine-tuning) is an interpolated result using recorded speckles. Specific hyperparameters and fine-tuning sample amount used in experiments are available in Ref. [83].

target. In these circumstances, the adaptive recursive algorithm always shows much better GFP (17–25) than traditional algorithm results (8–17). As seen from Fig. 4(c), with adaptive recursive algorithms, the enhancement achieved in GFP over the traditional one is 43%–108%. Considering the influence from the environment, the enhancement percentage achieved in experiments not being as high as that in simulation is reasonable, but significant improvements in focusing performance are demonstrated by both simulation and experiment. The dotted lines in Fig. 4(c) indicate the trend that with the increase of  $SUM_M$ , i.e., longer nonstationary processes, the enhancement percentage realized by the adaptive recursive algorithm keeps rising, suggesting that the merits of the adaptive recursive algorithm are becoming notable as the nonstationary processes are becoming longer. With the traditional fine-tuning

algorithm, when fine-tuning is conducted and when the original trained model is obtained ( $t = 0$ ), the difference between the medium's status increases overall with  $SUM_M$ . The low statistical correlation between these two statuses induces degradation in the modeling accuracy, increasing the difficulty of focusing recovery. Moreover, sending all of the samples altogether to fine-tune the original model suggests that the whole process is regarded as a stationary variation by the network, which, actually, may be nonstationary and consists of multiple stationary stochastic sub-processes.

The reduction percentage in the GTE achieved by the adaptive recursive algorithm over traditional one is shown in Fig. 4(d), reaching 30%–57%, indicating that the adaptive recursive algorithm is much better than the traditional method in terms of focusing tracking, which is consistent with the

simulation results. Moreover, as  $SUM_M$  becomes larger, the focusing tracking performance of the traditional algorithm keeps being exacerbated due to the lack of ability in timely recovery. By contrast, the adaptive recursive algorithm conducts fine-tuning successively during the whole nonstationary process; thus, the reduction percentage achieved by the adaptive recursive algorithm increases as  $SUM_M$  becomes larger, as suggested by the dotted lines in Fig. 4(d).

In situations where environmental disturbance occurs, the experimental results are shown in Figs. 5(d)–5(f). The sudden perturbation can be regarded as a stationary stochastic sub-process, whose SDT is very small, and the time duration is extremely short. Although the occurrence of perturbations leads to inevitable increase in the GTE, as observed in both Figs. 4(b) and 2(l), the adaptive recursive algorithm still demonstrates much lower GTE in experiments (2–15) than that achieved by the traditional algorithm (13–25). The reduction percentage in the GTE realized by the adaptive recursive method over the traditional one is 38%–93%. As for the GFP, the adaptive recursive algorithm achieves larger values all the time, ranging from 17 to 27, while values obtained by the traditional method are much smaller, being from 4 to 13. The enhancement percentage of the adaptive recursive algorithm over the traditional performance is 56%–444%, achieving significant improvement in the focusing performance, which has also been indicated in simulation. These results suggest that the advantage of recursive fine-tuning becomes more outstanding in this situation, as timely tracking is more and more important in long nonstationary processes. Meanwhile, the merits of adaptive adjustments of hyperparameters also become more notable since the requirement of fewer fine-tuning samples expedites the focusing recovery and leads to better recovery performance as well. In addition, Fig. 5(e) demonstrates that once a slower changing sub-process is detected ( $SDT = 400$  s), the adaptive recursive algorithm is able to make up the PBR loss caused by perturbations, which, again, is consistent with the simulation, as illustrated in Fig. 2(i).

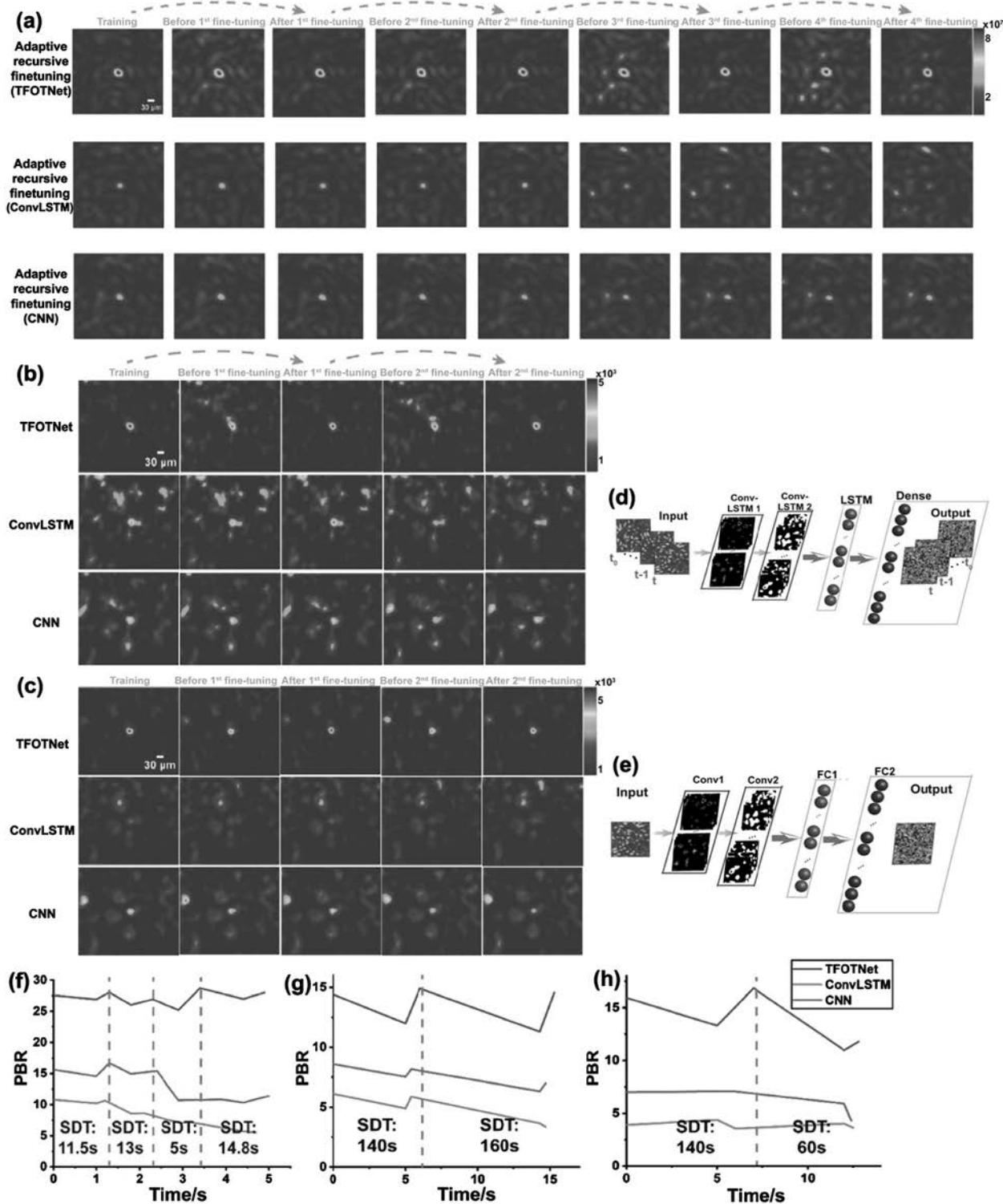
Actually, all of the enhancement and reduction results obtained in the six experiments agree well with the simulations. As seen from Figs. 4(a) and 4(c), with the increase of  $SUM_M$ , the GFP of the adaptive recursive method keeps rising; meanwhile, the enhancement percentage it can achieve over the traditional performance also enlarges, regardless of the occurrence of disturbance. In all circumstances, the adaptive recursive algorithm always demonstrates the best results. As for the GTE, the reduction percentage over the traditional algorithm also increases with  $SUM_M$ , no matter whether perturbations take place or not. These results suggest that the proposed adaptive framework is robust and even more promisingly attractive when the nonstationary process lasts longer or significant sudden PBR degradation occurs.

The speckle images recorded in the nonstationary processes indicated by Fig. 5(f) with the adaptive recursive and traditional algorithms are shown in Fig. 5(g). All speckle images are interpolated to  $253 \times 253$  for a better view, and the interpolation algorithm is based on splines. The diameter of the initial focused speckle is  $\sim 30 \mu\text{m}$ . Speckle patterns before and after each fine-tuning are demonstrated. As seen in Fig. 5(g),

the adaptive recursive algorithm can recover the focal point in time, and then the focus can remain over time. By contrast, with the traditional algorithm, due to the lack of ability of timely tracking, it cannot recover the focal point even though fine-tuning is conducted. It is worth noting that although we only report light focusing to a single position, the trained TFOTNet is capable of focusing light to an arbitrary position or multiple positions simultaneously on the image plane. As indicated above, during experiments, only the speckles before and after fine-tuning are recorded with the traditional algorithm, and the middle image in the bottom row in Fig. 5(g) is an interpolated result using recorded speckles.

### 3. Comparison of Light Focusing and Refocusing Performance with TFOTNet, Conventional ConvLSTM, and CNN

The ability of the conventional single-input-single-output ConvLSTM network [85,86] and single-input-single-output CNN in light focusing and refocusing through nonstationary scattering media is investigated using both simulation and experiments, and results are shown in Fig. 6. The structures of the conventional ConvLSTM and CNN are shown in Figs. 6(d) and 6(e), respectively. The conventional ConvLSTM network consists of two ConvLSTM layers, one LSTM layer, and one TimeDistributed dense layer working as the output layer. The input of the network is speckle patterns, while the output is their corresponding SLM patterns. All layers share the same parameters with their corresponding ones in TFOTNet, including kernel size, number of filters, activation function, etc. As for CNN, it consists of two convolutional layers, one fully connected layer, and the other fully connected layer serving as the output layer. Except timestep, which is not included in CNN, all of the other parameters are the same as the ConvLSTM network. With simulation, in the first step, the same 10,000 samples are used to train TFOTNet and CNN in order to obtain a focal point, and training results are shown in Fig. 6(a). All figures in Fig. 6(a) use the same colormap. The PBRs of the focused speckle achieved by TFOTNet and CNN are 41.5 and 15.6, respectively. As for the ConvLSTM network, 15,000 samples are used, which is an increase of 50% compared with that needed by TFOTNet. Nonetheless, the PBR of the focused speckle obtained with ConvLSTM is only 10.79, much lower than that achieved with the pre-trained TFOTNet (41.5). This phenomenon actually indicates a drawback of conventional ConvLSTM networks, where, as both temporal and spatial weights have to be learned during training, a large amount of samples are required. However, TFOTNet significantly enhances the modeling efficiency and effectively overcomes this drawback. After obtaining a focal point, during the fine-tuning phase with a nonstationary process, the same adaptive hyperparameters and fine-tuning samples are offered to TFOTNet, ConvLSTM, and CNN (except timestep). As seen in Figs. 6(a) and 6(f), with the same nonstationary process and fine-tuning algorithm, TFOTNet always exhibits the best performance in light focusing and refocusing. Nevertheless, with the ConvLSTM network or CNN, over time, the background becomes so bright that a single focal point is no longer able to be recovered, even though recursive fine-tuning is conducted. As for experimental results, 10,000 samples are sent to



**Fig. 6.** Comparisons about fine-tuning ability in a nonstationary process using three different networks (see Visualization 1). (a) Simulation results. Light focusing and refocusing performance recorded at different times through a nonstationary process using the adaptive recursive algorithm with TFOTNet (the first row), ConvLSTM (the second row), and CNN (the third row) is shown. All images use the same colormap and scale, and the color bars indicate the light intensity in arbitrary units. (b) Experimental results. Light focusing and refocusing results using three networks with the adaptive recursive algorithm in the same nonstationary process are shown. All speckle images use the same colormap and scale and are interpolated to  $253 \times 253$ . The color bars indicate the detected light intensity in arbitrary units. (c) Experimental results in the situation of low SNR. All speckle images use the same colormap and scale and are interpolated to  $253 \times 253$ . The color bars indicate the detected light intensity in arbitrary units. (d) Structure of a conventional single-input-single-output ConvLSTM. (e) Structure of a conventional single-input-single-output CNN. (f) Details of the nonstationary process and PBR with the three networks in simulation (a). (g) Details of the nonstationary process and PBR with the three networks in experiment (b). (h) Details of the nonstationary process and PBR with the three networks in experiment (c). (f)–(h) use the same legend.

initialize and train TFOTNet, ConvLSTM, and CNN. With TFOTNet, after training, a focused speckle can be obtained while using the other two networks, clear background speckles are observed with PBR dropping to less than 60% of that achieved by TFOTNet, as shown in Figs. 6(b) and 6(g). With adaptive recursive fine-tuning, a focused speckle can always be retained using TFOTNet through a nonstationary scattering medium; by contrast, the focal point is submerged over time when ConvLSTM or CNN is used, which agrees well with simulation results. Interestingly, experimental results demonstrate that in situations of low SNR, as shown in Figs. 6(c) and 6(h), among the three networks, only the proposed TFOTNet is able to obtain focus after training, even though the same training samples and parameters are used. In the situations indicated by Figs. 6(a)–6(c), the SNR of the training results of TFOTNet is calculated as 14, 12, and 10, respectively, which is defined as the ratio of the mean value of the signal and the standard deviation of the noises [87,88]. As seen in Fig. 6(c), even if some fine-tuning samples are offered, the fine-tuned ConvLSTM or CNN still cannot focus light through a nonstationary scattering medium. As indicated by Figs. 6(a) and 6(b), the refocusing performance of conventional ConvLSTM or CNN degrades over time; thus, it can be deduced that under the circumstances of low SNR a focal point can hardly be obtained with these two networks. This phenomenon manifests that TFOTNet is more robust to noises than conventional single-input-single-output networks.

As mentioned above, the SLM used in the experiment limits the fine-tuning sample collection speed, which further restricts the allowed changing speed of the scattering medium. Nevertheless, it should be emphasized here that there is no fundamental limitation on the speed of the proposed adaptive deep learning framework, since much faster modulators can be employed. With a current commercially available DMD whose frame rate has reached 23 kHz [82], both sample collection speed and the SDT of the altering medium can be improved nearly 4000 times. Thus, for the experimental results shown in Fig. 5, the SDT can be shortened to 3.5–113 ms, indicating that the proposed framework can be potentially applied for wavefront shaping in dynamic media, for instance, optical focusing and imaging at depths *in vivo* that alters at the speed of milliseconds [84]. Therefore, the proposed framework opens up a potential pathway to resolve the high demand of wavefront shaping on responding time, taking a significant step towards practical realizations. The proposed adaptive recursive fine-tuning approach applying millisecond variation media *in vivo* will be further studied and will be reported elsewhere.

#### 4. Comparison of Time Cost in Focusing Recovery with Various Algorithms

It should be noted that time cost in focusing recovery by the adaptive recursive fine-tuning algorithm and the two representative conventional wavefront shaping techniques, continuous sequential algorithm (CSA) and transmission matrix measurement, is discussed herein for comparison. Assuming a nonstationary process whose duration is  $t$ , on average, the medium status changes per  $\Delta t$ . For the adaptive recursive algorithm, on average,  $M$  samples are needed for each fine-tuning, and

totally  $M_{\text{total}}$  samples are used during the whole nonstationary course. In comparison, if CSA or transmission matrix measurement is adopted to recover the focusing performance through the changed medium, the iterative optimization process or transmission matrix measurement has to be repeated from the beginning. Thus, the time cost should be  $(KN^2 \times \frac{M_{\text{total}}}{M})/F$  and  $(4N^2 \times \frac{M_{\text{total}}}{M})/F$ , respectively, where  $N$  is the dimension of the SLM pattern,  $K$  is the pixel gray level, and  $F$  is the frame rate of the SLM.  $\frac{M_{\text{total}}}{M}$  represents how many times fine-tuning has been done during the whole nonstationary process, and, each time fine-tuning is conducted, CSA or transmission matrix measurement will also be run once for focusing recovery. As for the adaptive recursive algorithm, the total time spent in one fine-tuning is  $\frac{M}{F} + pt_p$ . The fine-tuning time cost consists of two parts: sample collection time and computational time. The sample collection time is independent of the SLM dimension  $N$ ; instead, it is determined by the sample amount  $M$  and the SLM frame rate  $F$ , written as  $\frac{M}{F}$ . The computational time is the product of epoch number  $p$  and time cost per epoch  $t_p$ . Considering that during the fine-tuning, only the last layer of the pre-trained network, which has  $N^2$  neurons, is adjusted,  $t_p$  is the function of  $N^2$ , that is  $t_p = g(N^2) = \lambda(N)N^2$ . Besides  $N$ ,  $\lambda(N)$  is also influenced by various factors such as network structures, computational engine, and amount of fine-tuning samples. It should be noted that taking a powerful GPU will reduce  $t_p$ , thus reducing  $\lambda(N)$ . As an example, in our work, with the TFOTNet and reported computation platform (Acer Predator G9-792, 16 Gb RAM, and a GTX 980M GPU), in simulation,  $t_p$  is 0.38, 0.4, 0.45, 0.49, 0.54, and 0.61 s when  $N$  is set as 8, 16, 32, 64, 128, and 256, respectively. The fine-tuning sample amount used here is 1000, which is the largest sample amount used in the reported experiments, and the timestep and batch size are set as 2 and 64, respectively. It can be expected that  $t_p$  can be further reduced if the fine-tuning sample amount is smaller, or a more powerful computational unit is adopted.  $\lambda(N) = t_p/N^2$  is calculated to vary from  $9.2 \times 10^{-6}$  to  $5.9 \times 10^{-3}$ . In our work,  $N = 32$ , and  $\lambda(N)$  is calculated to be  $4.4 \times 10^{-4}$ . For an intuitive comparison, herein, time cost in one focusing recovery process using the adaptive recursive fine-tuning algorithm, CSA, and transmission matrix measurement is given below based on the setup reported in this article, with the SLM pattern size being  $32 \times 32$  and frame rate of LCoS-SLM as 60 Hz.  $K$  varies with different setups, which can be set as 8 [22], 191 [89], or other values, and here we adopt  $K = 32$  to be consistent with our experimental settings. Hence, more than 9 min is needed by CSA to complete an iterative optimization process. To measure a new transmission matrix that represents the changed medium status, nearly 70 s is required. As indicated by experimental results in Fig. 5, with the adaptive recursive algorithm, time spent in each fine-tuning varies from 6 to 170 s with the frame rate of the SLM being only  $\sim 6$  Hz. Since 60 Hz is used for optimization time estimation, for a fair comparison, the fine-tuning time should be reduced by 10 times, varying from 0.6 to 17 s. Therefore, the proposed adaptive fine-tuning algorithm can improve the speed by 32–910 times and 4–113 times against CSA and transmission matrix measurement, respectively. In addition, to measure

a transmission matrix, interference between the modulated light and a reference light is required, which significantly increases the system complexity and reduces the utilization efficiency of the SLM, considering that part of the SLM pixels work as the reference. As CSA optimizes each pixel independently, the detected intensity improvement at the output plane is small, which may lead to errors in phase selection, especially when the SNR is low [90].

One promising application of the proposed framework is encryption. Recently, learning-based optical encryption has been reported [91] with parameters of the trained model as the security keys, achieving high security. In their study, static diffusers are used. With our framework, rotating diffusers can be applied to create much more complex scattering conditions, which enhances system security as more parameters are required to precisely model the process. More importantly, the introduction of a fine-tuning engine makes the system robust to attack. If it is sensed that current security key is partially eavesdropped, the diffuser can be rotated to create a new scattering situation to disable the leaked key. Meanwhile, with our adaptive fine-tuning system, security keys that fit for the new setting can be obtained rapidly, preventing loss due to attack. Demonstration of the idea is underway and will be reported elsewhere.

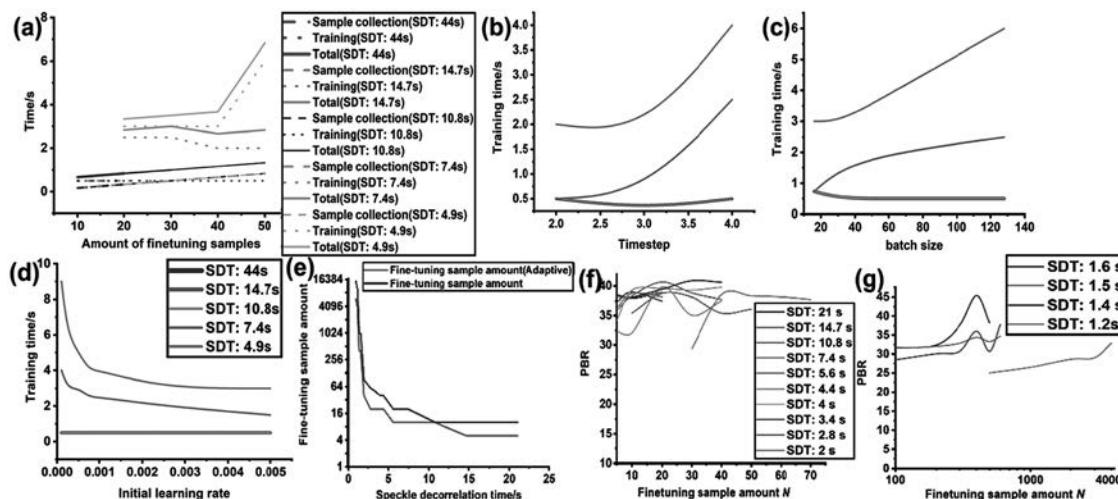
Lastly, but not the least, it should be admitted that since the fine-tuning cost is determined by the correlation between the medium statuses, it is natural that more samples and longer times are needed to recover a focal point in situations with a short SDT or dramatic disturbance. The proposed adaptive deep learning framework makes the best use of correlation between medium statuses to reduce the cost in focusing recovery as much as possible. However, in extreme cases where the medium decorrelates rapidly, or the persistence is low between iterations, a new training or optimization cycle may be required. The proposed adaptive recursive framework can achieve optimal tracking performance to a physical process change unless unpredictable.

#### 4. INFLUENCE OF HYPERPARAMETERS ON FINE-TUNING AND THE IMPLEMENTATION OF ADAPTIVE PBR TARGET

Hyperparameters including timestep, batch size, initial learning rate, and numbers of fine-tuning samples are investigated, respectively, as they will influence light refocusing performance or fine-tuning time cost. Simulation is conducted to evaluate the fine-tuning time cost to reach a pre-defined PBR target, as the above listed relevant parameters are individually varied with the medium changing at different speeds; then these results can be scaled in light of specific implementations. In all cases, fine-tuning is conducted with the same time interval after the original focal point is obtained. This is to ensure the same degree of medium change when different hyperparameters are investigated in a steadily altering situation, since the fine-tuning cost is directly related with the correlation amongst medium statuses. In this simulation, serving as an example, the PBR target is set as 37, and the time interval is set as 1 s. Nevertheless, values are not fixed; instead, they are adjustable according to specific setups. When one hyperpara-

meter is under test, all other hyperparameters remain as their default values, which are given in the inserted table in Fig. 2. From Figs. 7(a)–7(d), it can be seen that when medium altering is mild, and the SDT is larger than 10.8 s, varying one hyperparameter does not lead to significant differences in the fine-tuning time. With faster speed (SDT is smaller than 10.8 s), selecting suitable values for hyperparameters becomes more essential, as they are exerting growing influence on the fine-tuning time cost. Sample collection time is estimated using the maximal frame rate of commercial LCoS-SLM, which is generally 60 Hz. As for epochs, there is no doubt that more epochs require longer computation time and may lead to overfitting. On the other hand, more epochs may contribute to better light focusing performance.

Among all of the hyperparameters evaluated above, the amount of fine-tuning samples imposes the most significant influence on the fine-tuning time cost. A comparison of the required fine-tuning sample amount with and without adaptive adjustments of hyperparameters in the situations of different SDTs is shown in Fig. 7(e). As seen, without adaptive modifications of hyperparameters, the required sample amount is 2–3 times that needed by the adaptive algorithm to reach a pre-defined PBR target, resulting in much longer fine-tuning time cost. The result can be extended to other SDTs that are not tested here. During fine-tuning, if the number of newly collected samples is smaller than 100, then previously collected samples are included to concatenate with new samples; thus, a total of 1000 samples are used for fine-tuning to avoid overfitting. Increasing the fine-tuning sample amount theoretically leads to better focusing performance, which, however, also prolongs the fine-tuning process, as more time is spent at sample collection, indicating larger change to the medium status during this time period, which will degrade the fine-tuned PBR. With the goal of balancing the trade-off between the overall PBR and the fine-tuning time cost, we explore the relationship between PBR after fine-tuning and fine-tuning samples amounts using the adaptive algorithm when the medium changes at different speeds, with results shown in Figs. 7(f) and 7(g). In all cases, fine-tuning is conducted after a fixed time interval when the initial focused speckle is obtained, and it is chosen as 1 s as an example, which is scalable. With slow medium change (SDT larger than 2.8 s), less than 30 fine-tuning samples are sufficient to recover the PBR back to 37, which is regarded as an acceptable PBR threshold in this simulation. With faster medium changes, characterized by SDT ranging from 1.6 to 1.2 s, several hundred samples are needed to surpass the PBR threshold. With further expedition of the medium change (SDT lower than 1.2 s), up to several thousand samples are required. As longer time is required to collect more samples, the capability of tracking the medium change and keeping light focused will be affected. To mitigate this dilemma, an adaptive PBR target is employed; as a faster SDT accommodates a relatively lower PBR target, less time is needed. With instructive results shown in Fig. 7, for a certain SDT interval, the adaptive PBR target is defined as the mean value of the maximal and minimal PBR that can be achieved after fine-tuning, i.e.,  $PBR\ target = \frac{\max(PBR) + \min(PBR)}{2}$ , serving as a criterion to evaluate the focusing recovery performance.



**Fig. 7.** Influence of hyperparameters on fine-tuning time cost and performance when the scattering medium changes at various speeds. (a)–(d) The effect of sample amount, timestep, batch size, and initial learning rate on fine-tuning time cost under five circumstances where the scattering medium is changing at different speeds (indicated by lines of different colors and quantified by speckle decorrelation time). (e) The required amount of fine-tuning samples with and without adaptive adjustments of hyperparameters as the medium changes at different speeds. (f), (g) The relationship between the PBR after fine-tuning and the fine-tuning sample amount using the adaptive algorithm as the medium changes at different speeds. The default values of these hyperparameters used in the simulation are listed in Fig. 2.

Although different setups result in different initial focused speckles, the ratio between the PBR target and PBR of the initial focused pattern remains unchanged as long as the SDT is the same. By doing so, the presented results can be safely scaled to any other implementations.

## 5. CONCLUSION

In summary, the proposed deep learning-empowered adaptive wavefront framework is able to achieve focusing and fast refocusing of light through time-variant scattering media, which now allows a complex nonstationary stochastic process for the kind of first time, to the best of our knowledge. With the proposed adaptive recursive fine-tuning of TFOTNet, optical focusing can be recovered in time from degradation, which is much more rapid (with fundamental potential to achieve real-time) than both the traditional fine-tuning algorithm and the representative conventional methods, which require a new time- and/or resource-demanding optimization process. Simulation and experimental results agree very well and manifest the merits of the proposed framework. The experimental results indicate that with the proposed adaptive recursive framework, for all  $SUM_M$  investigated here, the GFP can be enhanced by 43%–444% against the traditional algorithm, and the GTE is reduced by 30%–93%. Moreover, as the nonstationary process is prolonged, both the GFP enhancement percentage and the GTE reduction percentage over the traditional algorithm increase. It can be expected that similar performance improvement can be achieved by other implementations of the proposed framework. All results shown in the article are scalable according to the specific implementations; with the proposed framework, light focusing can be safely retained in all realizations. As stated, with DMD and more powerful GPU, the proposed framework has potential to deal with scattering media of

SDT being several milliseconds; thus, it opens up a potential pathway to resolve the high demand of wavefront shaping on responding time, taking a significant step towards practical realizations.

**Funding.** Agency for Science, Technology and Research (A18A7b0058); National Natural Science Foundation of China (81627805, 81671726, 81930048); Guangdong Science and Technology Commission (2019A1515011374, 2019BT02X105); Hong Kong Innovation and Technology Commission (GHP/043/19SZ, GHP/044/19GD, ITS/022/18); Hong Kong Research Grant Council (25204416, R5029-19); Shenzhen Science and Technology Innovation Commission (JCYJ20170818104421564).

**Disclosures.** The authors declare no conflicts of interest.

<sup>†</sup>These authors contributed equally to this paper.

## REFERENCES

1. J. Bertolotti, E. G. Van Putten, C. Blum, A. Lagendijk, W. L. Vos, and A. P. Mosk, “Non-invasive imaging through opaque scattering layers,” *Nature* **491**, 232–234 (2012).
2. P. Lai, L. Wang, J. W. Tay, and L. V. Wang, “Photoacoustically guided wavefront shaping for enhanced optical focusing in scattering media,” *Nat. Photonics* **9**, 126–132 (2015).
3. Z. Yu, M. Xia, H. Li, T. Zhong, F. Zhao, H. Deng, Z. Li, D. Li, D. Wang, and P. Lai, “Implementation of digital optical phase conjugation with embedded calibration and phase rectification,” *Sci. Rep.* **9**, 1537 (2019).
4. J. Yang, Y. Shen, Y. Liu, A. S. Hemphill, and L. V. Wang, “Focusing light through scattering media by polarization modulation based generalized digital optical phase conjugation,” *Appl. Phys. Lett.* **111**, 201108 (2017).

5. J. Yang, J. Li, S. He, and L. V. Wang, "Angular-spectrum modeling of focusing light inside scattering media by optical phase conjugation," *Optica* **6**, 250–256 (2019).
6. Y. Shen, Y. Liu, C. Ma, and L. V. Wang, "Sub-Nyquist sampling boosts targeted light transport through opaque scattering media," *Optica* **4**, 97–102 (2017).
7. I. M. Vellekoop and A. P. Mosk, "Focusing coherent light through opaque strongly scattering media," *Opt. Lett.* **32**, 2309–2311 (2007).
8. S. M. Popoff, G. Lerosey, R. Carminati, M. Fink, A. C. Boccara, and S. Gigan, "Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media," *Phys. Rev. Lett.* **104**, 100601 (2010).
9. O. Katz, E. Small, and Y. Silberberg, "Looking around corners and through thin turbid layers in real time with scattered incoherent light," *Nat. Photonics* **6**, 549–553 (2012).
10. H. Yu, T. R. Hillman, W. Choi, J. O. Lee, M. S. Feld, R. R. Dasari, and Y. Park, "Measuring large optical transmission matrices of disordered media," *Phys. Rev. Lett.* **111**, 153902 (2013).
11. T. Chaigne, O. Katz, A. C. Boccara, M. Fink, E. Bossy, and S. Gigan, "Controlling light in scattering media non-invasively using the photo-acoustic transmission matrix," *Nat. Photonics* **8**, 58–64 (2014).
12. A. Sanjeev, Y. Kapellner, N. Shabairou, E. Gur, M. Sivnani, and Z. Zalevsky, "Non-invasive imaging through scattering medium by using a reverse response wavefront shaping technique," *Sci. Rep.* **9**, 12275 (2019).
13. A. Drémeau, A. Liutkus, D. Martina, O. Katz, C. Schülke, F. Krzakala, S. Gigan, and L. Daudet, "Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques," *Opt. Express* **23**, 11898–11911 (2015).
14. K. T. Takasaki and J. W. Fleischer, "Phase-space measurement for depth-resolved memory-effect imaging," *Opt. Express* **22**, 31426–31433 (2014).
15. O. Katz, E. Small, Y. Guan, and Y. Silberberg, "Noninvasive nonlinear focusing and imaging through strongly scattering turbid layers," *Optica* **1**, 170–174 (2014).
16. E. Edrei and G. Scarcelli, "Memory-effect based deconvolution microscopy for super-resolution imaging through scattering media," *Sci. Rep.* **6**, 33558 (2016).
17. X. Xu, H. Liu, and L. V. Wang, "Time-reversed ultrasonically encoded optical focusing into scattering media," *Nat. Photonics* **5**, 154–157 (2011).
18. B. Judkewitz, Y. M. Wang, R. Horstmeyer, A. Mathy, and C. Yang, "Speckle-scale focusing in the diffusive regime with time-reversal of variance-encoded light (TROVE)," *Nat. Photonics* **7**, 300–305 (2013).
19. S. Resisi, Y. Viernik, S. M. Popoff, and Y. Bromberg, "Wavefront shaping in multimode fibers by transmission matrix engineering," *APL Photon.* **5**, 036103 (2020).
20. X. Wei, Y. Shen, J. C. Jing, A. S. Hemphill, C. Yang, S. Xu, Z. Yang, and L. V. Wang, "Real-time frequency-encoded spatiotemporal focusing through scattering media using a programmable 2D ultrafine optical frequency comb," *Sci. Adv.* **6**, eaay1192 (2020).
21. G. Huang, D. Wu, J. Luo, Y. Huang, and Y. Shen, "Retrieving the optical transmission matrix of a multimode fiber using the extended Kalman filter," *Opt. Express* **28**, 9487–9500 (2020).
22. J. Thompson, B. Hokr, and V. Yakovlev, "Optimization of focusing through scattering media using the continuous sequential algorithm," *J. Mod. Opt.* **63**, 80–84 (2016).
23. D. B. Conkey, A. N. Brown, A. M. Caravaca-Aguirre, and R. Piestun, "Genetic algorithm optimization for focusing through turbid media in noisy environments," *Opt. Express* **20**, 4840–4849 (2012).
24. J. Luo, Z. Wu, D. Wu, Z. Liu, X. Wei, Y. Shen, and Z. Li, "Efficient glare suppression with Hadamard-encoding-algorithm-based wavefront shaping," *Opt. Lett.* **44**, 4067–4070 (2019).
25. Z. Wu, J. Luo, Y. Feng, X. Guo, Y. Shen, and Z. Li, "Controlling 1550-nm light through a multimode fiber using a Hadamard encoding algorithm," *Opt. Express* **27**, 5570–5580 (2019).
26. O. Katz, P. Heidmann, M. Fink, and S. Gigan, "Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations," *Nat. Photonics* **8**, 784–790 (2014).
27. Y. Liu, C. Ma, Y. Shen, J. Shi, and L. V. Wang, "Focusing light inside dynamic scattering media with millisecond digital optical phase conjugation," *Optica* **4**, 280–288 (2017).
28. D. Wang, E. H. Zhou, J. Brake, H. Ruan, M. Jang, and C. Yang, "Focusing through dynamic tissue with millisecond digital optical phase conjugation," *Optica* **2**, 728–735 (2015).
29. M. Chen, H. Liu, Z. Liu, P. Lai, and S. Han, "Expansion of the FOV in speckle autocorrelation imaging by spatial filtering," *Opt. Lett.* **44**, 5997–6000 (2019).
30. J.-H. Park, Z. Yu, K. Lee, P. Lai, and Y. Park, "Perspective: wavefront shaping techniques for controlling multiple light scattering in biological tissues: toward *in vivo* applications," *APL Photon.* **3**, 100901 (2018).
31. Y. Luo, S. Yan, H. Li, P. Lai, and Y. Zheng, "Focusing light through scattering media by reinforced hybrid algorithms," *APL Photon.* **5**, 016109 (2020).
32. E. Bossy and S. Gigan, "Photoacoustics with coherent light," *Photoacoustics* **4**, 22–35 (2016).
33. Z. Li, Z. Yu, H. Hui, H. Li, T. Zhong, H. Liu, and P. Lai, "Edge enhancement through scattering media enabled by optical wavefront shaping," *Photon. Res.* **8**, 954–962 (2020).
34. Y. Shen, Y. Liu, C. Ma, and L. V. Wang, "Focusing light through scattering media by full-polarization digital optical phase conjugation," *Opt. Lett.* **41**, 1130–1133 (2016).
35. Y.-K. Xu, W.-T. Liu, E.-F. Zhang, Q. Li, H.-Y. Dai, and P.-X. Chen, "Is ghost imaging intrinsically more powerful against scattering?" *Opt. Express* **23**, 32993–33000 (2015).
36. E. Edrei and G. Scarcelli, "Optical imaging through dynamic turbid media using the Fourier-domain shower-curtain effect," *Optica* **3**, 71–74 (2016).
37. B. Hwang, T. Woo, and J.-H. Park, "Fast diffraction-limited image recovery through turbulence via subsampled bispectrum analysis," *Opt. Lett.* **44**, 5985–5988 (2019).
38. B. Blochet, L. Bourdieu, and S. Gigan, "Focusing light through dynamical samples using fast continuous wavefront optimization," *Opt. Lett.* **42**, 4994–4997 (2017).
39. R. H. Benjamin Judkewitz, I. M. Vellekoop, I. N. Papadopoulos, and Y. Changhuei, "Translation correlations in anisotropically scattering media," *Nat. Phys.* **11**, 684–689 (2015).
40. J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Advances in Neural Information Processing Systems* (2012), pp. 341–349.
41. G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica* **6**, 921–943 (2019).
42. A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica* **4**, 1117–1125 (2017).
43. L. Waller and L. Tian, "Computational imaging: machine learning for 3D microscopy," *Nature* **523**, 416–417 (2015).
44. A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Phys. Rev. Lett.* **121**, 243902 (2018).
45. Y. Rivenson, Y. Wu, and A. Ozcan, "Deep learning in holography and coherent imaging," *Light Sci. Appl.* **8**, 1 (2019).
46. Y. Wu, Y. Rivenson, H. Wang, Y. Luo, E. Ben-David, L. A. Bentolila, C. Pritz, and A. Ozcan, "Three-dimensional virtual refocusing of fluorescence microscopy images using deep learning," *Nat. Methods* **16**, 1323–1331 (2019).
47. M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: a review," *IEEE Signal Process. Mag.* **34**, 85–95 (2017).
48. C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 295–307 (2015).
49. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* **521**, 436–444 (2015).
50. A. Turpin, I. Vishniakou, and J. D. Seelig, "Light scattering control in transmission and reflection with neural networks," *Opt. Express* **26**, 30911–30929 (2018).
51. Y. Zhang, C. Wu, Y. Song, K. Si, Y. Zheng, L. Hu, J. Chen, L. Tang, and W. Gong, "Machine learning based adaptive optics for doughnut-shaped beam," *Opt. Express* **27**, 16871–16881 (2019).

52. S. Cheng, H. Li, Y. Luo, Y. Zheng, and P. Lai, "Artificial intelligence-assisted light control and computational imaging through scattering media," *J. Innov. Opt. Health Sci.* **12**, 1930006 (2019).
53. Y. Li, Y. Xue, and L. Tian, "Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media," *Optica* **5**, 1181–1190 (2018).
54. S. Li, M. Deng, J. Lee, A. Sinha, and G. Barbastathis, "Imaging through glass diffusers using densely connected convolutional networks," *Optica* **5**, 803–813 (2018).
55. B. Rahmani, D. Loterie, G. Konstantinou, D. Psaltis, and C. Moser, "Multimode optical fiber transmission with a deep learning network," *Light Sci. Appl.* **7**, 69 (2018).
56. Y. Sun, J. Shi, L. Sun, J. Fan, and G. Zeng, "Image reconstruction through dynamic scattering media based on deep learning," *Opt. Express* **27**, 16032–16046 (2019).
57. Y. Wang, J. Zhang, H. Zhu, M. Long, J. Wang, and P. S. Yu, "Memory in memory: a predictive neural network for learning higher-order non-stationarity from spatiotemporal dynamics," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2019), pp. 9154–9162.
58. Q. Luo, J. A. Newman, and K. J. Webb, "Motion-based coherent optical imaging in heavily scattering random media," *Opt. Lett.* **44**, 2716–2719 (2019).
59. H. Yilmaz, E. G. van Putten, J. Bertolotti, A. Lagendijk, W. L. Vos, and A. P. Mosk, "Speckle correlation resolution enhancement of wide-field fluorescence imaging," *Optica* **2**, 424–429 (2015).
60. A. Porat, E. R. Andresen, H. Rigneault, D. Oron, S. Gigan, and O. Katz, "Widefield lensless imaging through a fiber bundle via speckle correlations," *Opt. Express* **24**, 16835–16855 (2016).
61. I. M. Vellekoop, "Controlling the propagation of light in disordered scattering media," *arXiv:0807.1087* (2008).
62. Z. Wei and X. Chen, "Deep-learning schemes for full-wave nonlinear inverse scattering problems," *IEEE Trans. Geosci. Remote Sens.* **57**, 1849–1860 (2018).
63. W. C. Chew and Y.-M. Wang, "Reconstruction of two-dimensional permittivity distribution using the distorted Born iterative method," *IEEE Trans. Med. Imaging* **9**, 218–225 (1990).
64. X. Chen, "Subspace-based optimization method for solving inverse-scattering problems," *IEEE Trans. Geosci. Remote Sens.* **48**, 42–49 (2009).
65. U. S. Kamilov and H. Mansour, "Learning optimal nonlinearities for iterative thresholding algorithms," *IEEE Signal Process. Lett.* **23**, 747–751 (2016).
66. K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.* **26**, 4509–4522 (2017).
67. E. Rueckert, M. Nakatenus, S. Tosatto, and J. Peters, "Learning inverse dynamics models in o(n) time with LSTM networks," in *IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)* (2017), pp. 811–816.
68. J. S. Marron and M. P. Wand, "Exact mean integrated squared error," *Ann. Stat.* **20**, 712–736 (1992).
69. S. Feng, C. Kane, P. A. Lee, and A. D. Stone, "Correlations and fluctuations of coherent wave transmission through disordered media," *Phys. Rev. Lett.* **61**, 834–837 (1988).
70. M. Breitkreiz and P. W. Brouwer, "Semiclassical theory of speckle correlations," *Phys. Rev. E* **88**, 062905 (2013).
71. P. Sebbah, *Waves and Imaging through Complex Media* (Springer, 2001).
72. L. Belfiore, A. Arkadan, and B. Lenhardt, "ANN inverse mapping technique applied to electromagnetic design," *IEEE Trans. Magn.* **37**, 3584–3587 (2001).
73. L. Li, J. Pan, W.-S. Lai, C. Gao, N. Sang, and M.-H. Yang, "Learning a discriminative prior for blind image deblurring," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), pp. 6616–6625.
74. J. Adler and O. Öktem, "Solving ill-posed inverse problems using iterative deep neural networks," *Inverse Prob.* **33**, 124007 (2017).
75. A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging: beyond analytical methods," *IEEE Signal Process. Mag.* **35**, 20–36 (2018).
76. J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, "One network to solve them all—solving linear inverse problems using deep projection models," in *Proceedings of the IEEE International Conference on Computer Vision* (2017), pp. 5888–5897.
77. J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Advances in Neural Information Processing Systems* (2014), pp. 3320–3328.
78. H.-C. Shin, H. R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, and R. M. Summers, "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imaging* **35**, 1285–1298 (2016).
79. L. Castrejon, Y. Aytar, C. Vondrick, H. Pirsiavash, and A. Torralba, "Learning aligned cross-modal representations from weakly aligned data," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 2940–2949.
80. V. Tran, S. K. Sahoo, and C. Dang, "Fast 3D movement of a laser focusing spot behind scattering media by utilizing optical memory effect and optical conjugate planes," *Sci. Rep.* **9**, 1 (2019).
81. M. M. Qureshi, J. Brake, H.-J. Jeon, H. Ruan, Y. Liu, A. M. Safi, T. J. Eom, C. Yang, and E. Chung, "*In vivo* study of optical speckle decorrelation time across depths in the mouse brain," *Biomed. Opt. Express* **8**, 4855–4864 (2017).
82. Z. Yu, H. Li, and P. Lai, "Wavefront shaping and its application to enhance photoacoustic imaging," *Appl. Sci.* **7**, 1320 (2017).
83. Y. Luo, S. Yan, H. Li, P. Lai, and Y. Zheng, "Datafile\_towards smart optical focusing," [https://drive.google.com/drive/folders/1\\_jbotdfKimgvNYVXqRMx5JykYpWXUZ?usp=sharing](https://drive.google.com/drive/folders/1_jbotdfKimgvNYVXqRMx5JykYpWXUZ?usp=sharing) (2020).
84. I. Nissilä, T. Noponen, J. Heino, T. Kajava, and T. Katila, "Diffuse optical imaging," in *Advances in Electromagnetic Fields in Living Systems* (Springer, 2005), Vol. 4.
85. A. Xavier, An Introduction to ConvLSTM, <https://medium.com/neuronio/an-introduction-to-convlstm-55c9025563a7> (2019).
86. S. Xingjian, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-C. Woo, "Convolutional LSTM network: a machine learning approach for precipitation nowcasting," in *Advances in Neural Information Processing Systems* (2015), pp. 802–810.
87. L. Kaufman, D. M. Kramer, L. E. Crooks, and D. A. Ortendahl, "Measuring signal-to-noise ratios in MR imaging," *Radiology* **173**, 265–267 (1989).
88. B. M. Welsh, "Speckle imaging signal-to-noise ratio performance as a function of frame integration time," *J. Opt. Soc. Am. A* **12**, 1364–1374 (1995).
89. J. W. Tay, P. Lai, Y. Suzuki, and L. V. Wang, "Ultrasonically encoded wavefront shaping for focusing into random media," *Sci. Rep.* **4**, 3918 (2014).
90. Z. Fayyaz, N. Mohammadian, and M. R. Avanaki, "Comparative assessment of five algorithms to control an SLM for focusing coherent light through scattering media," *Proc. SPIE* **10494**, 1049461 (2018).
91. L. Zhou, Y. Xiao, and W. Chen, "Learning complex scattering media for optical encryption," *Opt. Lett.* **45**, 5279–5282 (2020).

# PHOTONICS Research

PHOTONICS Research is devoted to fundamental and applied research progress in optics and photonics. All papers will be subject to peer review, including invited papers.

Submit manuscripts by using OSA's submission system on the website (see the "Authors" page at <http://prj.osa.org>). Submission is a representation that the manuscript has not been published previously and is not currently submitted for publication elsewhere. Transfer of copyright (online or hard-copy form) is necessary in order for the Publishers to continue disseminating research results as widely as possible.

**Manuscript Preparation:** Complete instructions are available on the "Authors" page at <http://prj.osa.org>.

All correspondence concerning submissions or peer review should be addressed to the Journal Coordinator at [prjournal@osa.org](mailto:prjournal@osa.org). Correspondence concerning papers in the production process should be addressed to the Managing Editor at [prjournal@siom.ac.cn](mailto:prjournal@siom.ac.cn). In all correspondence, reference should be made to the title, author, and manuscript number.

Copyright © 2021 Chinese Laser Press

## Editorial Office

Jie Ding (Publisher)

Chinese Laser Press,

No. 390, Qinghe Road, Jiading, Shanghai 201800, China

Kelly Cohen (Senior Publisher, OSA Journals)

OSA – The Optical Society

2010 Massachusetts Ave. NW, Washington, DC

20036-1023, USA

## Editor-in-Chief

Lan Yang, Washington University, St. Louis, USA

## Deputy Editors

Yu-Ao Chen, University of Science and Technology of China, China

Yuri Kivshar, Australian National University, Australia

## Associate Editors

Eric Cassan, C2N, Université Paris-Sud, France

Nunzio Cenno, University of Campania Luigi Vanvitelli, Italy

Giuseppe D'Aguanno, The Johns Hopkins University-Applied Physics Lab, USA

Chun-Hua Dong, University of Science and Technology of China, China

Yan Feng, Shanghai Institute of Optics and Fine Mechanics, CAS, China

Weibo Gao, Nanyang Technological University, Singapore

Frederic Gardes, University of Southampton, UK

Li Ge, City University of New York, USA

Junpeng Guo, University of Alabama in Huntsville, USA

Yong-Zhen Huang, Institute of Semiconductors, CAS, China

Xianmin Jin, Shanghai Jiao Tong University, China

Hao-Chung (Henry) Kuo, Taiwan Chiao Tung University, China

Xiaohang Li, King Abdullah University of Science & Technology, Saudi Arabia

Di Liang, Hewlett Packard Labs, USA

Jinyang Liang, INRS, Universite du Quebec, Canada

Zhiwen Liu, Pennsylvania State University, USA

Ryuji Morita, Hokkaido University, Japan

Cheng-Wei Qiu, National University of Singapore, Singapore

Stephan Reitzenstein, Technische Universität Berlin, Germany

Kebin Shi, Peking University, China

Zhipei Sun, Aalto University, Finland

Koji Yamada, AIST, Japan

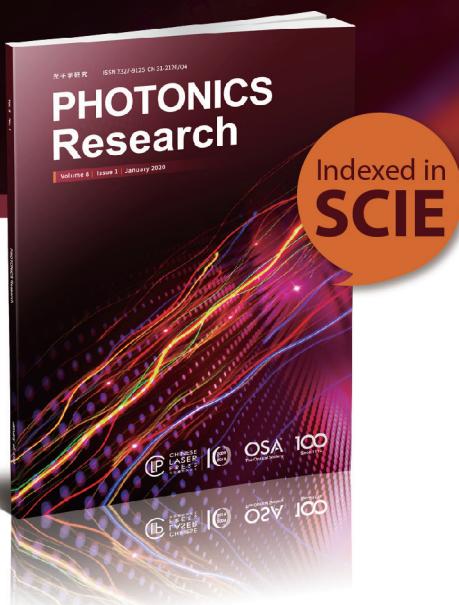
Jingbi You, Institute of Semiconductors, CAS, China

Han Zhang, Shenzhen University, China

Xinliang Zhang, Huazhong University of Science and Technology, China

# PHOTONICS Research

Disseminates fundamental and applied research progress in optics and photonics



7.080\*

2020 Impact Factor

\*According to the 2021 Journal Citation Reports®

## Author Benefits:

- Indexing in SCIE and Scopus
- Available on OSA Publishing's Digital Library
- Open access — Freely available to read, download and share
- High quality and productive peer review
- Professional editing and rapid publication
- Distinguished international editorial board
- Support from the editorial offices in Shanghai and Washington



Publish your next excellent paper in *Photonics Research*

For more information: E-mail: prjournal@osa.org; prjournal@siom.ac.cn Website: pr.researching.cn; prj.osa.org

