



## All-optical image transmission through a dynamically perturbed multimode fiber and a ring-core fiber using diffractive deep neural networks

JIANPING ZHU,<sup>1</sup> LEI ZHU,<sup>1,2,3,4,\*</sup> RAN GAO,<sup>1</sup> XISHUO WANG,<sup>5</sup> JIANXIN REN,<sup>6</sup> SHUAIDONG CHEN,<sup>7</sup> ZEXUAN JING,<sup>8</sup> ZHIPEI LI,<sup>1</sup> SHANTING HU,<sup>1</sup> QINGHUA TIAN,<sup>2,3,4</sup> BO TIAN,<sup>1</sup> HUAN CHANG,<sup>1</sup> DONG GUO,<sup>1</sup> XIAOLONG PAN,<sup>1</sup> ZHEYU WU,<sup>1</sup> XIANGJUN XIN,<sup>1</sup> AND BO LIU<sup>2,6,9</sup>

<sup>1</sup>School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China

<sup>2</sup>School of Electronic Engineering, BUPT, Beijing 100876, China

<sup>3</sup>State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>4</sup>Beijing Key Laboratory of Space-Ground Interconnection and Convergence, Beijing University of Posts and Telecommunications, Beijing 100876, China

<sup>5</sup>China Telecom Research Institute, State Key Laboratory of Optical Fiber and Cable Manufacture Technology, Beijing, China

<sup>6</sup>Institute of Optics and Electronics, Nanjing University of Information Science and Technology, Nanjing 210044, China

<sup>7</sup>Nanjing University of Posts and Telecommunications, Nanjing 210003, China

<sup>8</sup>School of Information and Communication Engineering, Beijing University of Posts and Communications, Beijing 100876, China

<sup>9</sup>Department of Electronic Engineering, Tsinghua University, Beijing 100084, China

\*7520220191@bit.edu.cn

Received 4 September 2024; revised 18 October 2024; accepted 18 October 2024; posted 21 October 2024; published 11 November 2024

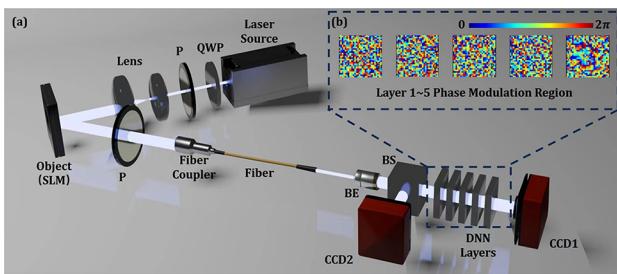
**In this study, we present an all-optical image reconstruction technique leveraging a diffractive deep neural network (D2NN) within a ring-core fiber (RCF) architecture. Orbital angular momentum (OAM) modes are employed to facilitate imaging transmission. We experimentally validate the efficacy of our approach for complex field diffractive image reconstruction through a multimode fiber (MMF) and RCF at a 1550 nm operating wavelength. The experimental results demonstrate the superior performance of the proposed method in mitigating RCF scattering-to-restoration transformation issues, significantly outperforming traditional MMF-based imaging correction techniques.** © 2024 Optica Publishing Group. All rights, including for text and data mining (TDM), Artificial Intelligence (AI) training, and similar technologies, are reserved.

<https://doi.org/10.1364/OL.541194>

In recent years, optical fiber imaging technology has garnered increasing attention. In these systems, when light propagates through strongly scattering media, the inherent randomness or non-uniformity of the medium induces the formation of speckle patterns at the output. Consequently, the underlying mechanisms of single-fiber imaging share similarities with those of imaging in scattering media, yet also exhibit distinct characteristics [1]. Particularly, multimode fiber (MMF)-based optical imaging has gained considerable attention over the past decade due to its ability to support numerous orthogonal spatial modes for image transmission. However, the performance of such systems is challenged by random phase delays and mode coupling

between different light modes, which induce scattering effects that degrade the overall imaging quality [2–4]. To mitigate the challenges associated with mode coupling, ring-core fibers (RCFs) supporting orbital angular momentum (OAM) mode groups have shown significant performance advantages. Unlike MMF, the weak coupling between distinct inter-mode groups in RCFs ensures that transmitted information remains less susceptible to distortion from environmental perturbations. Furthermore, the OAM-RCF imaging system enhances edge detail, which is critical for improving image clarity and precision. This configuration offers particular benefits for applications requiring high-quality imaging and phase-sensitive functionalities, including biomedical imaging and industrial nondestructive testing [5]. Despite these advancements, significant image distortion persists during fiber transmission under strong perturbations.

Recently, diffractive deep neural networks (D2NN) have emerged as a highly stable and robust framework, demonstrating significant potential across a wide range of optical applications [3,4]. This approach offers considerable advantages in terms of efficiency and precision, positioning D2NN as a promising tool for addressing complex challenges in the optical domain. D2NN represents a transformative shift in the landscape of machine learning [6]. By leveraging photons instead of electrons for computation, these systems perform operations as light propagates through the layers, involving both transmission and diffraction. D2NNs enable massive parallelism in optical signal processing, accommodating various wavelengths and polarization states [6–8]. Additionally, these networks can be physically realized, e.g., through 3-D printing, to address specific computational tasks. To date, limited research has focused on the



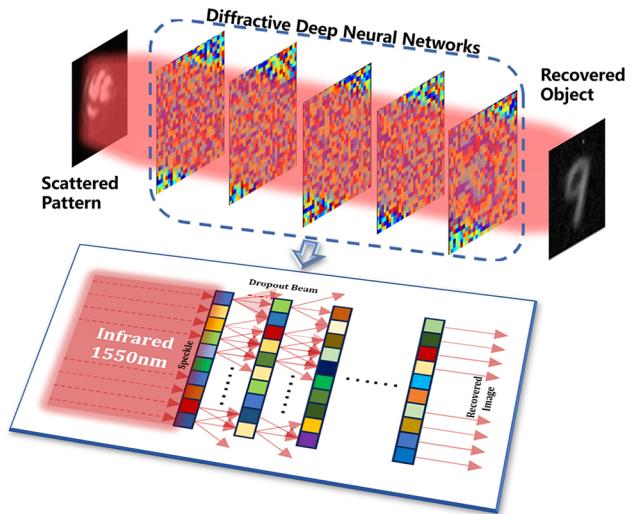
**Fig. 1.** (a) Schematic illustration of the experimental setup for the D2NN-aided fiber imaging in the MMF and RCF. QWP, quarter-wave plate; P, polarizer; SLM, spatial light modulator; BE, beam expander; BS, beam splitter; CCD, charge-coupled device. (b) Value of layers' phase modulation in the phase-modulation region on the board.

comprehensive exploration of all-optical image transmission through dynamically perturbed media using D2NN, particularly within the context of the RCF.

In this study, we introduce a novel method utilizing a D2NN as a reconstruction model to directly restore images distorted by transmission artifacts, specifically those arising from speckle patterns in the RCF. In contrast to the method proposed in [4,7], our approach demonstrates favorable reconstruction performance while employing a structure of reduced complexity. This simplification not only enhances the system's overall efficiency but also facilitates easier fabrication and implementation in real-world applications. To demonstrate the efficacy of the RCF-based approach, we compare its performance with that of the MMF under identical conditions. Our methodology employs the well-established Modified National Institute of Standards and Technology (MNIST) dataset as the input, subjecting it to different fiber transmissions. Our results affirm the successful reconstruction of the distorted speckle patterns back into the original MNIST images, thereby underscoring the potential of diffractive neural networks for mitigating optical field distortions and enhancing image reconstruction capabilities.

**Experiment system design.** As illustrated in Fig. 1(a), a schematic of the experimental RCF-D2NN and MMF-D2NN image transmission system are shown. The beam, after traversing a collimator lens, a quarter-wave plate (QWP) polarization system, and a magnifying lens system, is directed onto the spatial light modulator (SLM). The reflected beam, subsequently modified by an additional polarization stage, becomes an intensity-modulated image carrier, which is then coupled into the optical fiber. Both the MMF and RCF are selected as transmission mediums to demonstrate the robustness of our D2NN and its ability to recover distorted images from various types of fibers. Figure 1(b) illustrates the phase modulation across the layers of a five-layer D2NN used to reconstruct speckled images to their original digital form. The recovered images are recorded using a charge-coupled device (CCD1) camera. It is important to note that the image-speckle pairs used for training are captured under the same experimental conditions, with speckles being directly recorded by a separated CCD2.

The parameters of this system are described detailed as below: the laser source is a narrowband laser with an output wavelength of 1550 nm. The SLM used to display a digit object (PLUTO-2.1 Holoeye) consisted of  $1920 \times 1080$  microdisplays, each being 8  $\mu\text{m}$  in size. In the experiment, we merge the neighborhood units together as one bigger unit. The inner/outer radius



**Fig. 2.** Mechanism for scattering image reconstruction through D2NN layers.

of the core of the RCF was 4  $\mu\text{m}/8.5 \mu\text{m}$ , and the core and cladding diameters of GI-MMF were 50  $\mu\text{m}/125 \mu\text{m}$ . The CCD camera (Bobcat 320, Xeneth) utilized an in-house developed, temperature-stabilized InGaAs detector with a  $320 \times 256$  pixels resolution, offering frame rates of either 100 Hz or 400 Hz.

**Principles of operation for the proposed D2NN.** Here, we introduce an approach for optical image reconstruction directly from speckle patterns, as illustrated in Fig. 2, utilizing a modified optical neural network. As we all know, the D2NN operates entirely in the optical domain, requiring no reference or labeled data. The network is designed solely for classification tasks. Our proposed method enables the enhanced D2NN to reconstruct scattering images directly through an ultrafast all-optical process. As described by the D2NN framework, the image recovery process involves two key components: diffraction between layers and on-layer neuron operations. The diffraction aspect adheres to the Rayleigh–Sommerfeld diffraction equation, treating each neuron within a given layer as a secondary wave emitter. These neurons collectively shape the optical mode, as outlined in the subsequent discussion:

$$w_i^l = \frac{z - z_i}{r^2} \left( \frac{1}{2\pi r} + \frac{1}{j\lambda} \right) \exp \left( \frac{j2\pi r}{\lambda} \right), \quad (1)$$

where  $l$  represents the  $l$ -th layer,  $i$  means the modulation element in different location on the layer  $l$ ,  $\lambda$  is the wavelength of the illumination source,  $r = \sqrt{(x - x_i)^2 + (y - y_i)^2 + (z - z_i)^2}$ , and  $j = \sqrt{-1}$ . Let  $m_i^l = e^{j\phi}$  be the phase-modulation element in the  $l$ -th layer the  $i$ -th pixel element which delays the input wave for  $\Delta\phi$ . Theoretically, each neuron generates a secondary wave that diffracts in all directions, potentially influencing every neuron in the subsequent layer. However, given a specific spacing between consecutive layers, the wave strength of each neuron attenuates below the noise detection threshold after a certain propagation distance. This propagation distance in the next layer effectively defines the receptive field of the diffractive neural network. As a result, the output of the  $l$ -th layer can be expressed as a function of all the involved pixel elements:

$$\begin{aligned} \text{output}_i^l &= w_i^l \times \sum \text{output}_k^{l-1} \times m_i^l \\ &= w_i^l \times |A| e^{j\phi} \times e^{j\Delta\phi} = |A_w| e^{j\Delta\phi^*}. \end{aligned} \quad (2)$$

We summarize the propagation of light through a single diffractive layer in a matrix form as follows. The transmission process occurs in two stages: within the layer and between consecutive layers. Within the layer, both the phase and intensity are modulated, resulting in a modified light field. This modulated light then propagates through the air separating the layers of the network. By combining the entire transmission process across these stages, we obtain the following expression:

$$X^{out} = (M^n \odot \dots (M^2 \odot (M^1 \odot X^{in}) \cdot P^1) \cdot P^2 \dots) \cdot P^n, \quad (3)$$

where  $n$  represents the number of layers in the network. A single diffractive layer is insufficient to achieve the inference capability characteristic of a multi-layer D2NN architecture. The multi-layer design of a D2NN provides significant flexibility within a constrained physical volume, facilitating the training of complex transformation functions that map the input plane to the output plane.

**Training setup and processing.** During the parameterization training process, each signal and operator within the system is treated as a complex-valued matrix, encompassing the D2NN input image, on-layer phase modulation, and inter-layer propagation. These components are combined using distinct operations. The on-layer phase modulation is represented as a learnable parameter matrix, which is updated through gradient back-propagation, while the inter-layer propagation is modeled by a predefined matrix, governed by the physical distance between layers. It is important to note that all layers are assigned the same dimensions, and light propagation follows the principles of diffraction, as illustrated in Fig. 2. Additionally, the diffraction range is limited to nearby elements, leading to the occurrence of “dropout beams,” wherein some sub-beams at the margins no longer reach the subsequent layer.

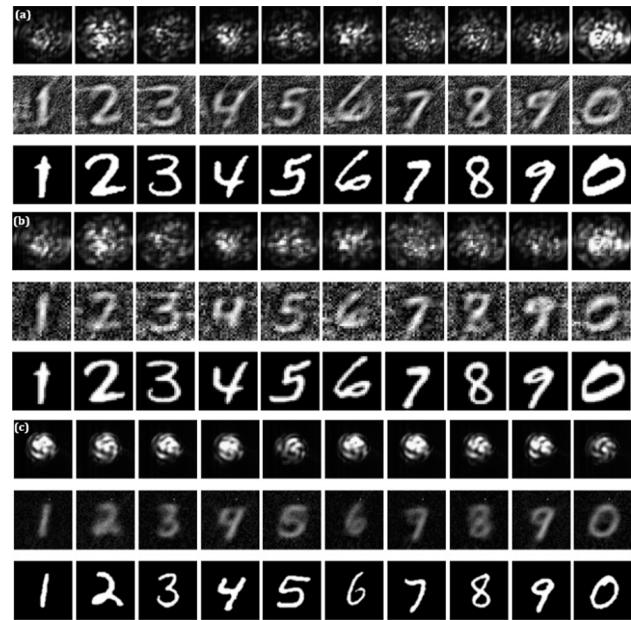
The training dataset comprises pairs of MNIST digit images and their corresponding fiber-transmitted speckle patterns (MMF/RCF) and down-sampled for resolution purposes. In this study, the layer size of the D2NN is matched to the resolution of the down-sampled speckle patterns. For model training, the dataset was divided into 9000 training samples and 1000 test samples, with a training period of 100 epochs. After training, the model was tested on previously unseen speckle patterns and their corresponding MNIST digits, which were excluded from the training process. The results demonstrate the D2NN model’s robust generalization capabilities for this specific task. The trained network successfully reconstructs the digits from previously scattering speckles, effectively extracting key features from the distortion experienced during image transmission through the fiber.

When employing the RCF dataset for training, validation, and testing in the D2NN image recovery, it was necessary to optimize the performance of the OAM beam transmission through the RCF. To achieve this, two spiral phase plates (SPP) were introduced: one positioned between P2 and the coupler and the other between BE and BS to modulate and demodulate the Gaussian and OAM beams. During transmission through the RCF, the  $l = +3$  OAM beam excited the  $l = -3$  state, leading to a superposition of the Gaussian and  $l = -6$  states in the subsequent SPP recovery process. Consequently, the D2NN training and processing involved handling two modal states simultaneously, even though the Gaussian light ( $l = 0$ ) contained the relevant or effectively usable information. The interference from the  $l = -6$  state hindered the D2NN training process, preventing it from accurately capturing the approximately linear mapping between

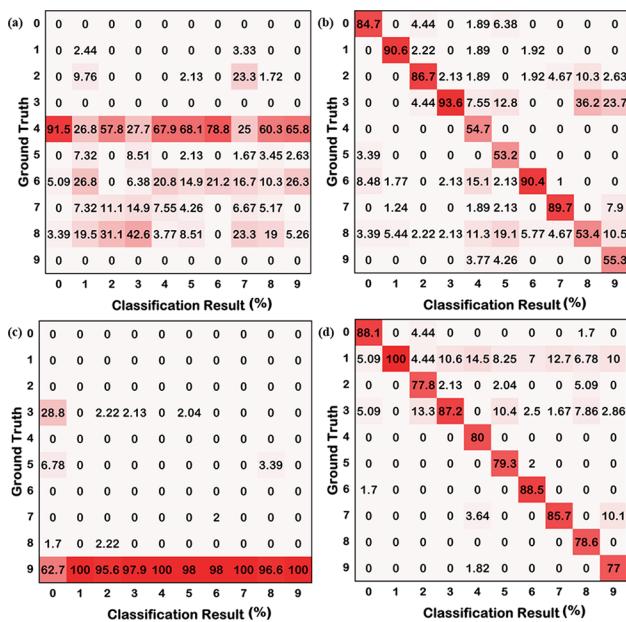
the light spot and the digit. As a result, the recovery of MNIST digit images suffered, particularly in terms of detail fidelity.

To overcome the superposition limits in RCF-OAM and successfully reconstruct objects within a complex scattering imaging scenario, we modified the loss function during the training process, replacing the cross-entropy error function with the Structural Similarity Index Measure (SSIM) loss function. SSIM serves as a representative index for assessing the similarity between the reconstructed images and the ground truth. To validate the effectiveness of the SSIM loss function, we performed a comparative analysis of the reconstruction performance under identical conditions, utilizing a conventional loss function based on the mean squared error (MSE) as a benchmark. This comparison enables a clearer evaluation of the advantages offered by the SSIM-based approach in enhancing the image recovery quality. The D2NN model was trained over a maximum of 100 epochs, with a batch size of 64, utilizing the Adam optimizer set to a learning rate of 0.001 to minimize the SSIM loss function. The network was implemented using a PyTorch 2.0.1 software and executed on an NVIDIA GeForce RTX 4090 graphics processing unit (GPU) and an Intel(R) Xeon(R) Platinum 8336C CPU @ 2.30 GHz (CPU). For an output size of  $64 \times 64$ , each training session takes approximately 40 min and each inference takes at the speed of light. It can even solve the ill-posed RCF speckles.

**Performance discussion and evaluation.** The test and reconstructed dataset acquired from fiber results are shown in Fig. 3. As the resolution increases from  $28 \times 28$  to  $64 \times 64$ , the recovered images exhibit clearer and more recognizable details. For instance, the lower curve of the digit “6” and the top of the digit “9” become more distinguishable, while the overall recovery performance improves for digits such as “4” and “8,” which display more complete structures derived from the original MNIST dataset. Notably, the parallelism of optical processing ensures that the increase in resolution primarily impacts the training time, without affecting the real-time speed of signal processing.



**Fig. 3.** Scattering speckles, reconstructed results, and the ground truth from “0” to “9” of held-out test dataset of (a) MMF, down-sampled to  $64 \times 64$ ; (b) MMF, down-sampled to  $28 \times 28$ ; and (c) RCF (ill-posed), down-sampled to  $64 \times 64$ .



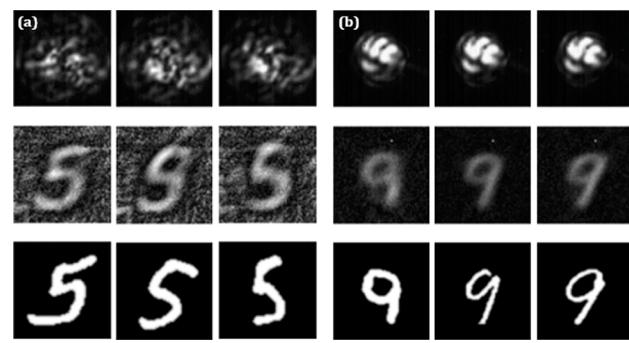
**Fig. 4.** Classification accuracy results between the MMF and RCF with/without D2NN reconstruction under down-sampled to  $64 \times 64$  (a) MMF, without D2NN; (b) MMF, with D2NN; (c) RCF, without D2NN; and (d) with D2NN (expressed in percentage form).

**Table 1. MSE and SSIM Index of the Results Recovered from MMF and RCF Speckles**

MMF	RCF	Percentage of Improvement
0.2554	0.1620	-36%
0.7050	0.7818	+11%

In the case of the RCF dataset, the higher concentration of light intensity during transmission leads to a corresponding concentration in the recovered intensity distribution, resulting in superior performance compared to the MMF. For a clearer comparison between the MMF and RCF, the classification results through a regular CNN with and without compensation are presented in Fig. 4. Comparison of MSE and SSIM (with a window size of 5) between the recovered images and the original images is presented in Table 1. To enhance the accuracy of performance evaluation, regions containing zero values are excluded prior to calculating the Structural Similarity Index Measure (SSIM) for the restored areas. This refinement ensures that the SSIM computation focuses solely on regions with meaningful image data, thereby providing a more representative assessment of the image restoration quality. The figure and table illustrate that although the recovered digits from the RCF appear dimmer due to the reduced overall luminance, the RCF outperforms the MMF. This is attributed to the stronger mode coupling in the MMF, whereas the RCF exhibits less leakage in OAM modes. Despite these challenges, our D2NN demonstrates impressive recovery of distorted images in both the MMF and RCF, showcasing its robustness across different transmission scenarios.

To assess the efficacy of the RCF-OAM D2NN across digits with varying structural features, we compare identical digits exhibiting different characteristics within both MMF and RCF setups, as illustrated in Fig. 5. This figure demonstrates the successful recovery of the same digits, such as “5” and “9,” presented with different writing styles and orientations. Table 1



**Fig. 5.** Scattering speckles and the reconstructed pattern of the same digit of (a) “5” in the MMF, down-sampled to  $64 \times 64$ , and (b) “9” in the RCF (ill-posed), down-sampled to  $64 \times 64$ .

further confirms that the RCF-OAM reconstruction exhibits adaptability under the SSIM condition. While the MMF reconstruction demonstrates superior performance under the MSE index, both MMF and RCF-OAM systems exhibit suboptimal retrieval parameters, a notable contrast to the improved outcomes observed when using the SSIM index. It is observed that the layers of the network do not merely learn to reconstruct a specific digit image but rather extract and generalize the distortive effects introduced by the fiber channel during transmission.

**Conclusion.** In this study, we introduce a method for reconstructing scattering images using a D2NN as a significant approach to addressing specialized fiber imaging tasks. We experimentally validate a novel integration of a modified D2NN within an OAM-based RCF imaging system, employing the SSIM loss function to mitigate mode superposition effects among different OAM modes. Our experimental results highlight the efficacy of the proposed method, particularly when compared to MMF systems across various loss functions. This approach not only offers ultra-fast processing speeds but also holds promise for fundamental applications in rapid distortion compensation in diverse fiber imaging systems.

**Funding.** National Key Research and Development Program of China (2022YFB2903104); National Natural Science Foundation of China (62225503, 62305027, 62105026, 62306038, 62205023, 62205151); Open Beijing Natural Science Foundation (JQ24018); Open Fund of IPOC (BUPT) (IPOC2022A04, IPOC2024ZZ01); BUPT Innovation and Entrepreneurship Support Program (2024-YCA052).

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

**Data availability.** Data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

**Supplemental document.** See Supplement 1 for supporting content.

## REFERENCES

1. A. K. Ghatak and K. Thyagarajan, *An Introduction to Fiber Optics* (Cambridge University Press, 1998).
2. C. Lutzweiler and D. Razansky, *Sensors* **13**, 7345 (2013).
3. J. Hu, D. Mengu, D. C. Tzarakis, et al., *Nat. Commun.* **15**, 1525 (2024).
4. Y. Zhang, Q. Zhang, H. Yu, et al., *Sci. Adv.* **10**, eadn2205 (2024).
5. Z. Wu, R. Gao, S. Zhou, et al., *Laser Photonics Rev.* **18**, 2300624 (2024).
6. X. Lin, Y. Rivenson, N. T. Yardimci, et al., *Science* **361**, 1004 (2018).
7. E. Goi, S. Schoenhardt, and M. Gu, *Nat. Commun.* **13**, 7531 (2022).
8. M. S. Sakib Rahman and A. Ozcan, *ACS Photonics* **8**, 3375 (2021).