

Neural network based design of metagratings

Sandeep Inampudi, and Hossein Mosallaei

Citation: [Appl. Phys. Lett.](#) **112**, 241102 (2018); doi: 10.1063/1.5033327

View online: <https://doi.org/10.1063/1.5033327>

View Table of Contents: <http://aip.scitation.org/toc/apl/112/24>

Published by the [American Institute of Physics](#)

Articles you may be interested in

[Electric-field induced second-harmonic generation of femtosecond pulses in atmospheric air](#)
Applied Physics Letters **112**, 241101 (2018); 10.1063/1.5030171

[Strain-induced modulation of near-field radiative transfer](#)
Applied Physics Letters **112**, 241104 (2018); 10.1063/1.5037468

[Plasmonic nanoparticle lithography: Fast resist-free laser technique for large-scale sub-50 nm hole array fabrication](#)
Applied Physics Letters **112**, 223101 (2018); 10.1063/1.5025096

[High-efficiency generation of Bessel beams with transmissive metasurfaces](#)
Applied Physics Letters **112**, 191901 (2018); 10.1063/1.5023553

[Transmodal elastic metasurface for broad angle total mode conversion](#)
Applied Physics Letters **112**, 241905 (2018); 10.1063/1.5032157

[Structured ultrasound microscopy](#)
Applied Physics Letters **112**, 251901 (2018); 10.1063/1.5026863

AIP | Conference Proceedings

**Get 30% off all
print proceedings!**

Enter Promotion Code **PDF30** at checkout



Neural network based design of metagratings

Sandeep Inampudi and Hossein Mosallaei

Electrical and Computer Engineering Department, Northeastern University, 360 Huntington Ave., Boston, Massachusetts 02115, USA

(Received 4 April 2018; accepted 16 May 2018; published online 13 June 2018)

Metagratings are flat and thin surfaces that rely on unique, periodically repeating (non-gradient), arbitrary shaped light scattering units for wave manipulation. **However, the absence of an empirical relationship between the structural and diffraction properties of the units enforces utilization of brute force numerical optimization techniques to determine the unit shape for a desired application. Here, we present an artificial neural network based methodology to develop a fast-paced numerical relationship between the two.** We demonstrate the training and the performance of a numerical function, utilizing simulated diffraction efficiencies of a large set of units, that can instantaneously mimic the optical response of any other arbitrary shaped unit of the same class. We validate the performance of the trained neural network on a previously unseen set of test samples and discuss the statistical significance. We then utilize the virtually instantaneous network operations to inverse design the metagrating unit shapes for a desired diffraction efficiency distribution. The proposed inter-disciplinary combination of advanced information processing techniques with Maxwell's equation solvers opens a pathway for the fast-paced *prediction* of metagrating designs rather than full wave computation. *Published by AIP Publishing.* <https://doi.org/10.1063/1.5033327>

Metasurfaces^{1–8} are new age optical elements that demonstrated several unconventional forms of light manipulating devices using flat and ultra-thin patterned layers. However, the typical metasurface design procedure of arranging subwavelength sized resonant light scatterers guided by the required phase gradient confronts with achievable phase resolution and fabrication limitations.^{9–12} To surpass these limitations, a new class of ultrathin, flat optical wave manipulation layers known as metagratings^{12–18} has been recently evolved to demonstrate that similar degree of wave engineering could be achieved by properly designing the geometrical and optical properties of unique, uniformly repeating, wavelength order sized units. For example, a bianisotropic unit has been demonstrated with unitary reflection efficiency with a minimal fabrication challenge,¹² a free form shaped unit has been demonstrated for large angle beam scanning performance^{13,17} and coupled dielectric waveguides have been demonstrated for a broadband beam scanning and splitting applications.¹⁸

Despite their great potential, the pace of development of metagratings in comparison to conventional metasurfaces is not up to the demand because of the absence of a clear design strategy and an analytical, numerical, or empirical relationship between the shape of the unit and the far-field intensity pattern. Determination of unit shape for a particular application often relies on computationally expensive and time consuming numerical evolutionary optimization techniques.^{13,17}

On the other hand, in a virtually unrelated field of information technology, new strategies of computational algorithms have been under huge development in the form of machine learning¹⁹ and deep learning²⁰ techniques that stem from artificial neural network analysis.²¹ Neural network algorithms are developed based on a heuristic vision to imitate the functionality of a biological brain where by processing the information about many examples through several “artificial neurons,” inferences are made on the relation

between the input and output set of parameters. The basic idea of neural network algorithms is to *predict* the solution to a problem by learning from a predefined set of examples rather than solving the problem directly. Neural network algorithms in the context of optics have been previously utilized for microscopy,²² antenna design,²³ lens-less imaging,²⁴ multi-layered thin films,²⁵ multi-layered nano-particles,²⁶ and H-shaped light scatterers.²⁷ Here, we utilize the neural network prediction algorithms to develop a numerical relationship between the shape of a metagrating unit and its diffraction efficiency (DE).

The basic unit of a neural network is a neuron (also known as a node or unit). A neural network typically consists of anything from a few tens to even millions of interconnected nodes distributed in a sequence of layers, as shown in Fig. 1. The number of nodes in the first and last layers is equal to the number of input and output parameters of the problem, respectively. Each node receives input from some other nodes, or from an external source (called bias b) and computes an output. Each input has an associated weight (w), which is assigned on the basis of its relative importance to other inputs. The node applies a function f (known as transfer function) to the weighted sum of its inputs, as shown in Fig. 1. The objective of the training process is to compute the weights and biases of each node by utilizing a training dataset. The weights and biases are computed using iterative optimization techniques, such as gradient descent optimization technique, to minimize the prediction error on the training set (details in [supplementary material](#)).

To demonstrate the concept of utilizing artificial neural networks in the context of metagrating design, we consider light scattering from an arbitrary shaped periodic dielectric structure placed on a reflecting surface with a spacer layer, as shown in Figs. 2(a) and 2(b). The metagrating, under plane wave illumination in the normal direction, scatters

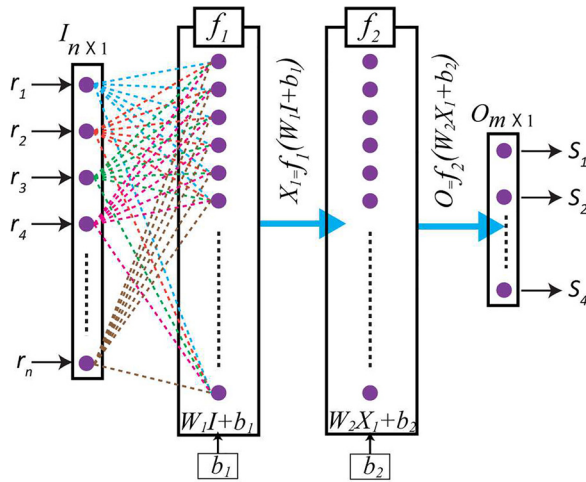


FIG. 1. A schematic representation of the architecture of a neural network.

light into diffraction orders along symmetric angles defined by the periodicity of the structure; however, the diffraction efficiency (DE) in each direction could be asymmetric based on the shape of the building block. Computing the DE into each angle for an arbitrary shaped unit usually involves solving Maxwell's equations with full wave solvers such as finite difference time domain technique (FDTD),²⁸ finite element method (FEM)²⁹ or rigorous coupled wave analysis (RCWA).³⁰ The inverse problem of predicting the shape of the building block for a desired set of DE in each angle is an extremely tedious task involving the computation of thousands of shapes using numerical optimization techniques such as a genetic algorithm.^{13,17,31–33} For each desired target set of DEs, the optimization methods re-search the whole parameter space consuming immense amount of time and computational resources.

In the case of neural networks, we compute the DEs of thousands of arbitrary shaped units “once” for the purpose to train the network function. Once trained, the network will quickly mimic the DE patterns of a gigantic set of arbitrary shaped building blocks using a few matrix operations. In the training set, for computational convenience, the shape of each unit is parameterized as a sixteen sided polygon with sixteen vertices whose positions could be varied arbitrary within the periodic bounds. Assuming each vertex in polar coordinates as $[(r_i, \theta_i), i = 1-16]$, the polar angles (θ_i) of the vertices are uniformly distributed between 0 and 2π . Therefore, the shape of the unit is now completely defined only by the radius coordinates (r_i) of the vertices. Note that the 16-dimensional space of radii vector with at least 10 steps along each dimension requires a dataset of 10^{16} arbitrary shapes for an accurate training process which is far beyond the practical limits of computational resources and time. In order to ensure that a smaller set of samples will incur the qualities of all possible shapes of the 16-dimensional space inside the work plane (unit cell), we consider a training dataset comprising 8 categories, as shown in Fig. 2(c). In the 8 categories, the radii are defined such that each 5000 set of shapes is restricted between the limits $(0 < r_i < 0.25)$, $(0.25 < r_i < 0.5)$, $(0.50 < r_i < 0.75)$, $(0.75 < r_i < 1.0)$, $(0.0 < r_i < 0.50)$, and $(0.50 < r_i < 1.0)$. We then forced additional 10000 shapes to have four-fold symmetry with radii between $(0.0 < r_i < 1.0)$, and another additional 50000 shapes are taken as purely random values between $(0.0 < r_i < 1.0)$. The radii of all the shapes can now be represented as a matrix “X” with 16 rows and (a total of) 90000 columns.

The chosen periodicity of the metagrating and the wavelength of incident light will result in a total of 13 diffraction orders inside the propagation regime.³⁴ The aim of the artificial neural network function (N_{net}) is to relate the 16 radii

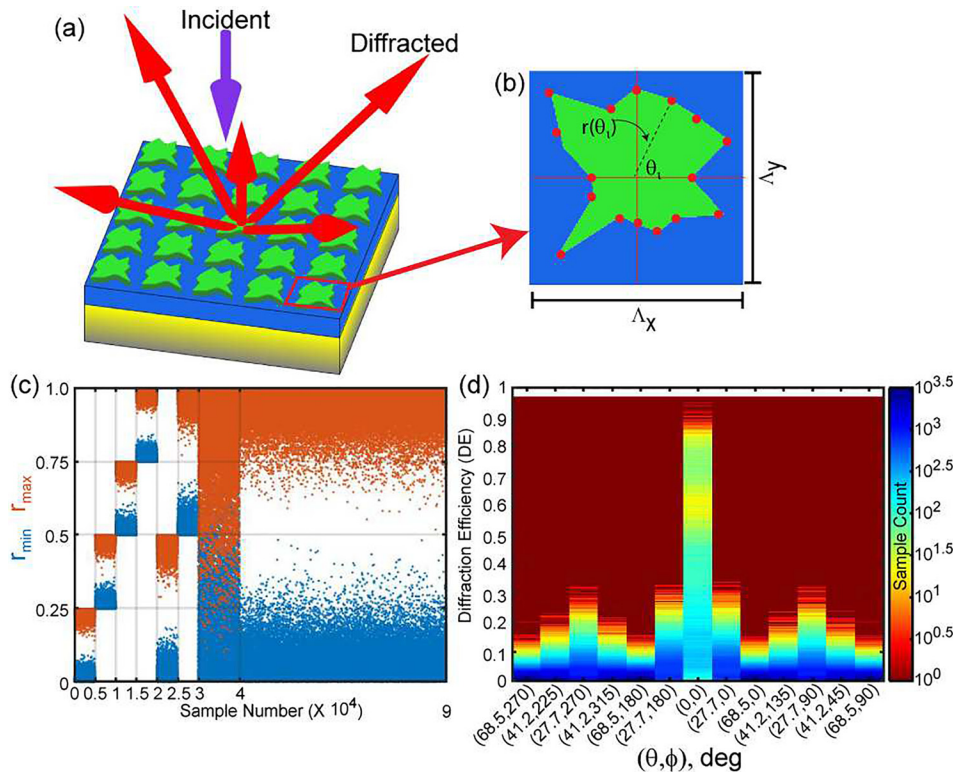


FIG. 2. (a) A schematic of a metagrating with an arbitrary shaped unit. (b) A representative shape of the unit using a 16-sided polygon parametrized using 16 radii values of the vertices. The unit dimensions are $\Lambda_x = \Lambda_y = 2.15\lambda_0$, where $\lambda_0 = 1.55 \mu\text{m}$. The spacer layer and the grating are made of silicon ($\epsilon = 12.67$) with heights of 400 nm each. The back mirror is assumed as gold. (c) The distribution of the range of radii values into 8 categories. (d) Histogram distribution of computed DEs of 13 resulting diffraction orders of 90000 units. The angle pairs on the x-axis represent the angle of the diffraction order in spherical coordinates (θ, ϕ) .

values (r_i , $i = 1-16$) of the shape to the DEs of the resulting 13 diffraction orders (D_j , $j = 1-13$). Mathematically, the relationship can be represented as

$$D = N_{net}(r), \quad (1)$$

where “ r ” and “ D ” represent the column vectors of radii and DEs, respectively.

The computation of DEs is carried out using an in-house built RCWA solver.^{30,35} The RCWA algorithm is computationally inexpensive and is customized to run in parallel utilizing the parallel computing toolbox of MATLAB. **On a workstation with 24 cores, the computation of DEs of 90 000 shapes consumed around 50 h.**

The computed dataset of DEs of the 13 propagation diffraction orders of 90 000 samples can be represented as a matrix “ Y ” with 13 rows and 90 000 columns. The histogram distribution of computed DEs of each diffraction order of all shapes is shown in Fig. 2(d). The responsibility of the neural network is now to compute the set of weight and bias matrices that represent a generalized numerical relationship between the matrices “ X ” and “ Y ”. Here, we trained the neural network using training packages based on feedforward network layers³⁶ available in MATLAB. **After many test runs, a feedforward network architecture is chosen with 3 layers (1-input, 1 hidden, and 1-output) with 16, 250 and 13 nodes, respectively** (details in [supplementary material](#)).

During the training process, the 90 000 sample shapes are further randomly divided into three subsets of 63 000 (70%), 13 500 (15%), and 13 500 (15%) for training, validation and testing purposes, respectively. The training dataset is utilized to iteratively compute the optimized values of weights and biases, while the validation dataset aids in fine-tuning the network parameters and determining further gradient directions. In other words, the training set directly participates in computing the weights and biases but the validation set occasionally participates in avoiding over-fitting of the data. Finally, the test dataset is used to perform an unbiased analysis of trained neural network characteristics such as mean square error (MSE) and regression. Each iteration of the training process to minimize the prediction error is known as an epoch (details in [supplementary material](#)). The MSEs of the training, validation and test datasets at each epoch are presented in Fig. 3(a). The training is considered to have reached an optimal position when the validation and test errors are relatively low and nearly equal to each other. Such an optimum position is reached after 100 epochs for our current dataset. Figure 3(b) presents the histogram of error distribution (between the actual and predicted DEs using the optimized weights and biases) of the datasets. The observed fact that the train, test and validation datasets follow the same error distribution demonstrates that the neural network is a good fit of the data without over-fitting or under-fitting.

After successful training of the function (N_{net}), we further tested the performance of the function for canonical shape geometries (like squares and circles that are deliberately avoided from the training set). Figures 3(c) and 3(d) represent the comparison between the DEs of light computed directly using RCWA (solid lines) and the trained network

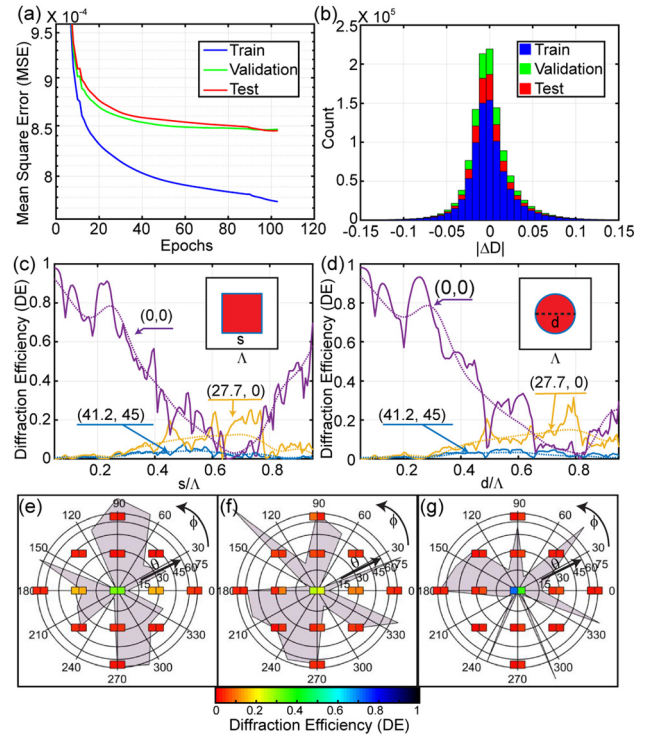


FIG. 3. Validation and error analysis of the neural network. (a) Mean square error of components of the training process at each epoch. (b) Error histogram distribution of the optimum trained network around 100th epoch. Comparison of DEs of three diffraction orders along the marked (θ, ϕ) directions computed using RCWA (solid lines) and predicted using the N_{net} function (dotted lines) for (c) square and (d) disc shaped unit cells as a function of their shape parameters. The computed and predicted DE values of diffraction orders for a few representative cases with an average error of 0.007 (e), 0.02 (f) and 0.04 (g) are presented for comparison. Double square markers are used to represent the diffraction angle (position in the θ, ϕ -plane), the computed DE values (left box color), and the predicted DE value (right box color) simultaneously.

(dotted lines). The network successfully reproduced the trends of DEs for all the diffraction angles. We opine that accurate prediction of each and every oscillation of the DE will be possible in future with the development of advanced neural network architectures. Figure 3(e)–3(g) presents the comparison of DE of a few representative arbitrary shaped samples with the average error distribution around 0.007, 0.02, and 0.04, respectively.

The fast-paced prediction of the DEs of arbitrary shapes make the (N_{net}) function a very appropriate tool to perform the inverse design of the unit for a specific target set of DEs. Here, we aim at retrieving the shape parameters (the radii) of the unit for a desired set of DEs (D_d) using an interior-point optimization algorithm.³⁷ The optimization process aims at determining the optimal set of radii (r) that satisfy the condition

$$\min \frac{1}{13} \sum_{j=1}^{13} |N_{net,j}(r_i) - D_{d,j}|^2 \text{ subject to } \{0 < r_i < 1\}. \quad (2)$$

First, to test the performance of the N_{net} function in the inverse design process, we required a full specular reflection of light in the normal direction along $[(\theta, \phi) = (0^\circ, 0^\circ)]$. The results are shown in Fig. 4(a). As expected, the N_{net} function

predicted the radii (r) of all 16 vertices to be negligible, which minimizes the light coupling to diffraction orders. (The back mirror reflects the light completely in the normal direction.) Figures 4(b) and 4(c) presents the predicted metasurating shapes for a typical beam scanning application of conventional metasurfaces where the normal incident light is required to predominately couple to one of the diffraction orders (at oblique angles).^{1–7} The resultant shapes in these cases are complex for a quantitative discussion; however, gradients and inclinations can be qualitatively observed in the shapes. Figure 4(d)–4(f) present the predicted shapes for multifunctional metasurface applications,^{33,38} where the incident light is distributed into two desired on-axis and off-axis symmetric angles (diffraction orders), respectively. The predicted shape for the uniform distribution of DE along on-axis symmetric angles [Fig. 4(d)] is (as expected) a one-dimensional diffraction grating. Similarly, the predicted shape for the uniform distribution of DE along off-axis symmetric angles [Fig. 4(e)] is nearly a one-dimensional diffraction grating rotated to 45° . The predicted shape for the uniform distribution of DE along asymmetric

angles [Fig. 4(f)] is again found to be complex for a qualitative discussion.

Note that one can perform the optimization process using full wave Maxwell's equation solvers by searching around for the best suitable candidate for the dataset for a target set of DEs; however, the search can be dramatically accelerated if the Maxwell's equation solver is replaced with the neural network function. The total computational time to predict all the six designs shown in Fig. 4 using the trained N_{net} function is less than a minute where the full Maxwell's equation solvers might require several hours.

Note that the N_{net} function is only trained for one configuration of the spacer and grating heights. Hence, the possible range of DE of a given diffraction order is limited to the range of values in the dataset. For example, as shown in Fig. 2(c), the maximum DE of the diffraction order along $(\theta, \phi) = (27.7^\circ, 180^\circ)$ is of the order of 0.5. Hence, the N_{net} function cannot predict any design with DE greater than 0.5 for the particular angle. This limitation is reflected in its lower efficiency of the predicted designs to diffract light to higher angles in comparison to a conventional graded pattern

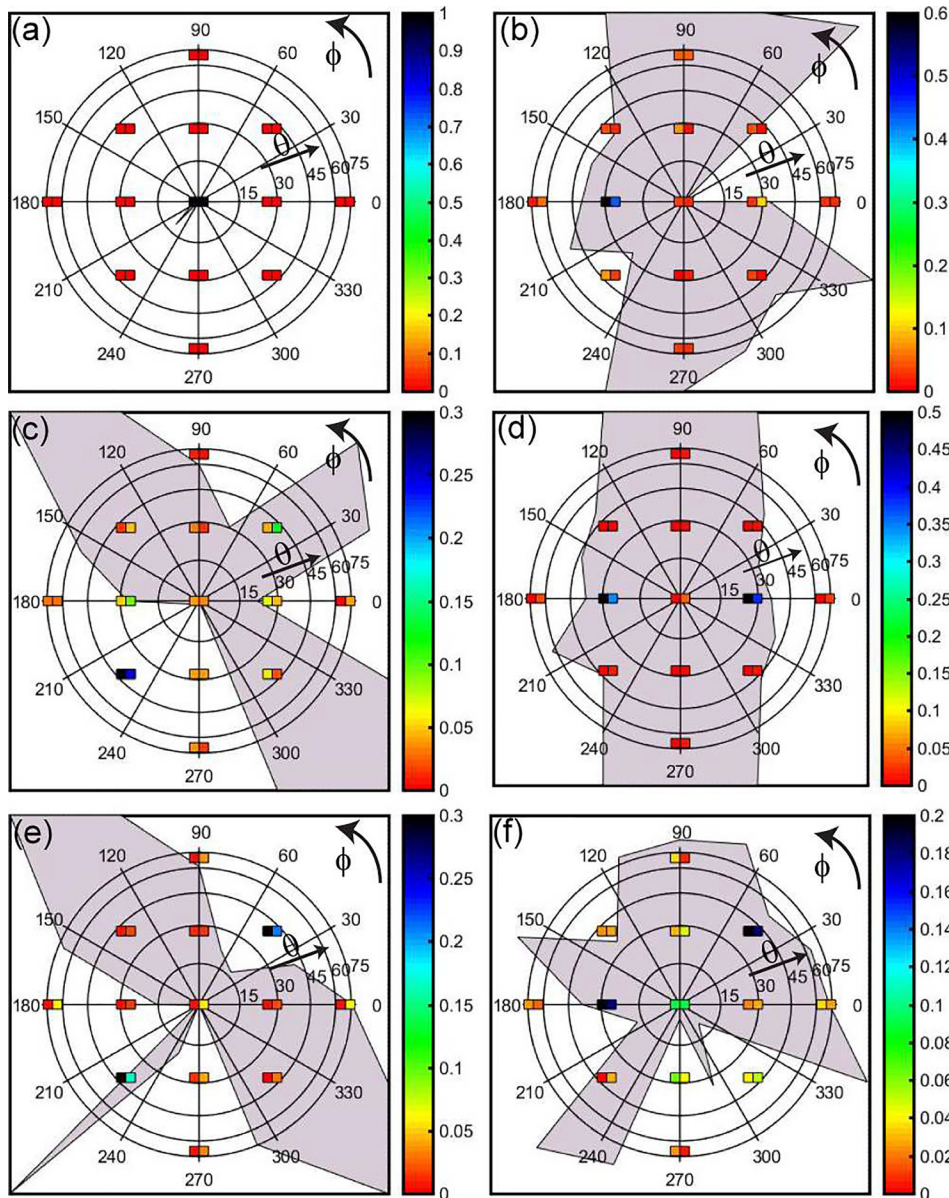


FIG. 4. Inverse prediction of unit shapes. (a)–(f) The DEs and units predicted for six choices of light distribution. Double square markers are used in each panel to represent the diffraction angle (position), the aimed DE (left box color) and the actual DE (right box color). The background of each panel represents the predicted unit shape. Panel (a) represents reflection in the normal direction $(\theta, \phi) \approx (0^\circ, 0^\circ)$, (b) represents the dominant reflection into a single on-axis angle $(\theta, \phi) = (27.7^\circ, 180^\circ)$, (c) represents dominant reflection into a single off-axis diffraction angle $(\theta, \phi) \approx (41.2^\circ, 225^\circ)$, (d) represents uniform reflection into two on-axis symmetric angles $[(\theta, \phi) = (27.7^\circ, 0^\circ)$ and $(\theta, \phi) = (27.7^\circ, 180^\circ)]$, (e) represents uniform reflection into two off-axis symmetric angles $[(\theta, \phi) \approx (41.2^\circ, 225^\circ)$ and $(\theta, \phi) \approx (41.2^\circ, 45^\circ)]$ and (f) represents uniform scattering into two asymmetric angles $[(\theta, \phi) \approx (27.7^\circ, 180^\circ)$ and $(\theta, \phi) \approx (41.2^\circ, 45^\circ)]$.

metasurface. Addition of the layer height as one of the parameters of training could enhance the performance of the current metagrating.

Another limitation for utilization of the N_{net} function is the uncertainty in the prediction of material loss. The presence of scatterers on the plasmonic back mirror inevitably results in some absorption loss. Hence, the sum of DEs of all diffraction orders does not always equal to 1 as we expect in an ideal zero-loss case. The sum of DEs of all diffraction orders of the 90 000 samples utilized during the training process fall in the range between 0.6 and 1 (0%–40% absorption loss) with nearly equal probability of all values in between. Hence, an average of 20% loss is assumed in all prediction cases in Fig. 4 [except panel (a)]. The surplus DE (1-assumed absorption loss-sum of required DEs of the diffraction orders) is randomly distributed into the remaining angles during inverse design optimization. The above discussed disadvantages of the neural network based design of metagrating could be attributed to limitations of both the prediction based neural network design methodology and the phase gradient less design.

To conclude, in this work, we design, train and test a neural network architecture to instantaneously mimic the optical response of metagratings with arbitrary unit shapes. We presented the capability of the neural networks for independently predicting the diffraction efficiency of a previously unseen large set of arbitrary shaped building blocks. We demonstrated that the network architecture also enables a fast-paced inverse design process of metagrating devices for a desired set of diffraction efficiencies. The proposed scheme allows the extension of the prediction mechanism to other physical and optical parameters of the host materials and a broad range of frequency spectra. We believe that the interdisciplinary combination of prediction algorithms with full wave solvers of Maxwell's equation opens a new pathway of on-the-fly device design processes for many nanophotonic applications that are particularly beneficial in the case of the evolving new forms of materials with dynamic control of optical^{39,40} and structural properties.⁴¹ The simplicity of the network function allows effortless integration of electromagnetic computations into larger Multi-physics modeling domains such as multicomponent optical networks, thermal radiation transport, quantum electrodynamics, digital holography, density functional theory, etc. We emphasize that the demonstrated inverse design process is just one immediate application example where the neural network is integrated within numerical optimization techniques.

See [supplementary material](#) for a detailed discussion on gradient descent technique for computing weights and biases of the network and building the neural network architecture.

This work was supported in part by the U.S. Air Force Office of Scientific Research (AFOSR) (No. FA9550-14-1-0349).

- ¹N. Yu, P. Genevet, M. A. Kats, F. Aieta, J.-P. Tetienne, F. Capasso, and Z. Gaburro, *Science* **334**, 333 (2011).
- ²X. Ni, N. K. Emani, A. V. Kildishev, A. Boltasseva, and V. M. Shalaei, *Science* **335**, 427 (2012).
- ³M. Farnahini-Farahani, J. Cheng, and H. Mosallaei, *J. Opt. Soc. Am. B* **30**, 2365 (2013).
- ⁴H.-T. Chen, A. J. Taylor, and N. Yu, *Rep. Prog. Phys.* **79**, 076401 (2016).
- ⁵S. B. Glybovski, S. A. Tretyakov, P. A. Belov, Y. S. Kivshar, and C. R. Simovski, *Phys. Rep.* **634**, 1 (2016).
- ⁶P. Genevet and F. Capasso, *Rep. Prog. Phys.* **78**, 024401 (2015).
- ⁷P. Genevet, F. Capasso, F. Aieta, M. Khorasaninejad, and R. Devlin, *Optica* **4**, 139 (2017).
- ⁸A. Forouzmmand and H. Mosallaei, *Adv. Opt. Mater.* **5**, 1700147 (2017).
- ⁹N. M. Estakhri and A. Alù, *Phys. Rev. X* **6**, 041008 (2016).
- ¹⁰V. Asadchy, M. Albooyeh, S. Tsvetkova, A. Díaz-Rubio, Y. Ra'di, and S. Tretyakov, *Phys. Rev. B* **94**, 075142 (2016).
- ¹¹A. Epstein and G. V. Eleftheriades, *J. Opt. Soc. Am. B* **33**, A31 (2016).
- ¹²Y. Ra'di, D. L. Sounas, and A. Alu, *Phys. Rev. Lett.* **119**, 067404 (2017).
- ¹³D. Sell, J. Yang, S. Doshay, R. Yang, and J. A. Fan, *Nano Lett.* **17**, 3752 (2017).
- ¹⁴A. Pors, M. G. Nielsen, and S. I. Bozhevolnyi, *Optica* **2**, 716 (2015).
- ¹⁵A. Epstein and O. Rabinovich, *Phys. Rev. Appl.* **8**, 054037 (2017).
- ¹⁶Y. Ra'di and A. Alu, *ACS Photonics* **5**, 1779 (2018).
- ¹⁷J. Yang, D. Sell, and J. A. Fan, *Ann. Phys.* **530**, 1700302 (2018).
- ¹⁸M. Khorasaninejad and F. Capasso, *Nano Lett.* **15**, 6709 (2015).
- ¹⁹I. H. Witten, E. Frank, M. A. Hall, and C. J. Pal, *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, 2016).
- ²⁰Y. LeCun, Y. Bengio, and G. Hinton, *Nature* **521**, 436 (2015).
- ²¹S. Haykin, *Neural Networks: A Comprehensive Foundation*, 2nd ed. (Prentice Hall PTR, Upper Saddle River, NJ, USA, 1998).
- ²²Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, and A. Ozcan, *Optica* **4**, 1437 (2017).
- ²³B. Liu, H. Aliakbarian, Z. Ma, G. A. Vandenbosch, G. Gielen, and P. Excell, *IEEE Trans. Antennas Propag.* **62**, 7 (2014).
- ²⁴A. Sinha, J. Lee, S. Li, and G. Barbastathis, *Optica* **4**, 1117 (2017).
- ²⁵D. Liu, Y. Tan, E. Khoram, and Z. Yu, *ACS Photonics* **5**, 1365 (2018).
- ²⁶J. Peurifoy, Y. Shen, L. Jing, Y. Yang, F. Cano-Renteria, B. Delacy, M. Tegmark, J. D. Joannopoulos, and M. Soljačić, *Proc. SPIE* **10526**, 1052607 (2018).
- ²⁷I. Malkiel, A. Nagler, M. Mrejen, U. Arieli, L. Wolf, and H. Suchowski, preprint [arXiv:1702.07949](#) (2017).
- ²⁸A. Taflov and S. C. Hagness, *Computational Electrodynamics: The Finite-Difference Time-Domain Method* (Artech House, 2005).
- ²⁹B. A. Szabo and I. Babuška, *Finite Element Analysis* (John Wiley & Sons, 1991).
- ³⁰M. Moharam, E. B. Grann, D. A. Pomet, and T. Gaylord, *J. Opt. Soc. Am. A* **12**, 1068 (1995).
- ³¹S. Inampudi, J. Cheng, and H. Mosallaei, *Appl. Opt.* **56**, 3132 (2017).
- ³²S. Inampudi, J. Cheng, M. M. Salary, and H. Mosallaei, *J. Opt. Soc. Am. B* **35**, 39 (2018).
- ³³J. Cheng, S. Inampudi, and H. Mosallaei, *Sci. Rep.* **7**, 12228 (2017).
- ³⁴M. Born and E. Wolf, *Principles of Optics: Electromagnetic Theory of Propagation, Interference and Diffraction of Light* (Elsevier, 2013).
- ³⁵C. M. Roberts, S. Inampudi, and V. A. Podolskiy, *Opt. Express* **23**, 2764 (2015).
- ³⁶K. Hornik, M. Stinchcombe, and H. White, *Neural Networks* **2**, 359 (1989).
- ³⁷R. A. Waltz, J. L. Morales, J. Nocedal, and D. Orban, *Math. Program.* **107**, 391 (2006).
- ³⁸S. M. Kamali, E. Arbabi, A. Arbabi, Y. Horie, M. Faraji-Dana, and A. Faraon, *Phys. Rev. X* **7**, 041056 (2017).
- ³⁹A. Forouzmmand and H. Mosallaei, *J. Opt.* **18**, 125003 (2016).
- ⁴⁰A. Forouzmmand, M. M. Salary, S. Inampudi, and H. Mosallaei, *Adv. Opt. Mater.* **6**, 1701275 (2018).
- ⁴¹W. K. Liu, E. G. Karpov, and H. S. Park, *Nano Mechanics and Materials: Theory, Multiscale Methods and Applications* (John Wiley & Sons, 2006).