

# Broadband achromatic metalens design based on deep neural networks

XIPENG AN,<sup>1,†</sup> YUE CAO,<sup>1,†</sup> YUNXUAN WEI,<sup>1</sup> ZHIHAO ZHOU,<sup>1</sup> TIE HU,<sup>1</sup> XING FENG,<sup>1</sup> GUANGQIANG HE,<sup>2</sup> MING ZHAO,<sup>1,3</sup> AND ZHENYU YANG<sup>1,\*</sup>

<sup>1</sup>School of Optical and Electronic Information, Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology, Wuhan 430074, China

<sup>2</sup>State Key Laboratory of Advanced Optical Communication Systems and Networks, Department of Electronic Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

<sup>3</sup>e-mail: zhaoming@hust.edu.cn

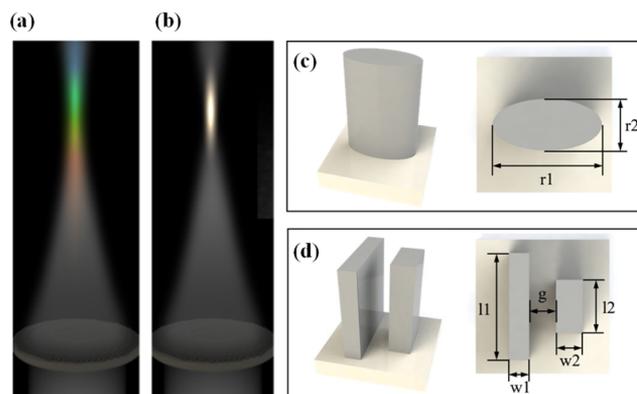
\*Corresponding author: zyang@hust.edu.cn

Received 13 April 2021; revised 19 June 2021; accepted 5 July 2021; posted 7 July 2021 (Doc. ID 427221); published 6 August 2021

For the design of achromatic metalenses, one key challenge is to accurately realize the wavelength dependent phase profile. Because of the demand of tremendous simulations, traditional methods are laborious and time consuming. Here, a novel deep neural network (DNN) is proposed and applied to the achromatic metalens design, which turns complex design processes into regression tasks through fitting the target phase curves. During training,  $x - y$  projection pairs are put forward to solve the phase jump problem, and some additional phase curves are manually generated to optimize the DNN performance. To demonstrate the validity of our DNN, two achromatic metalenses in the near-infrared region are designed and simulated. Their average focal length shifts are 2.6% and 1.7%, while their average relative focusing efficiencies reach 59.18% and 77.88%. © 2021 Optical Society of America

<https://doi.org/10.1364/OL.427221>

Metasurfaces [1], composed of two-dimensional subwavelength meta-unit arrays, have been widely studied because of their exceptional capacity in manipulating optical amplitude, phase, and polarization [2]. Metasurface-based lenses, also called metalenses [3–5], benefit imaging systems with high integration and complementary metal oxide semiconductor (CMOS) compatibility [6]. Previous works have demonstrated the excellent performance of monochromatic metalenses [7], but the realization of achromatic metalenses is still challenging. Metalenses' chromatic dispersions [8] arise from both material and light propagation, leading to wavelength dependent focal length shifts [Fig. 1(a)] and thus to the chromatic blur in imaging. To obtain the achromatic metalens shown in Fig. 1(b), one solution is to compensate for dispersion [9] with meta-units designed to simultaneously satisfy the wavelength and position dependent phase profile. Such meta-units are usually selected from large amounts of nanostructures with known electromagnetic (EM) responses that are traditionally calculated by numerical approaches such as the finite difference time domain (FDTD) method and finite element method (FEM). Some pioneering



**Fig. 1.** Illustrations of principles of achromatic metalenses and meta-units. (a) Schematic of a chromatic metalens. (b) Schematic of an achromatic metalens. (c) Geometric view of the elliptical meta-unit. The EM response is tuned by changing the major axis  $r_1$  and the minor axis  $r_2$ . (d) Geometric view of the nanofin meta-unit. The EM response is tuned by changing the length  $l_1$ ,  $l_2$  and width  $w_1$ ,  $w_2$  of two rectangular nanopillars and the gap  $g$ .

achromatic metalens designs have been proposed in both visible and near-infrared bands [10–12]; even a multifunctional metalens has been designed [13], but with limited efficiencies and flexibility because of the requirement of large numerical simulations.

Different from numerical approaches, deep learning [14,15] is a data-driven method, which has dramatically pushed forward the development of imaging processing, speech recognition, and natural language processing [16–18]. Moreover, deep learning has recently been introduced to nanophotonics [19–23].

Inspired by the extraordinary regression analysis ability of deep learning, we establish a deep neural network (DNN) to unscramble the relationship between meta-units and their EM responses for broadband achromatic metalens design. Furthermore, the target-fitting approach is applied to satisfy phase requirements of broadband achromatic meta-units. In detail, we overcome the phase jump problem with  $x - y$  projection pairs to improve the DNN training precision, and

we generate no-geometry data to optimize the network. To demonstrate the performance of the well-trained DNN, two near-infrared achromatic metalenses based on different kinds of datasets are designed and simulated.

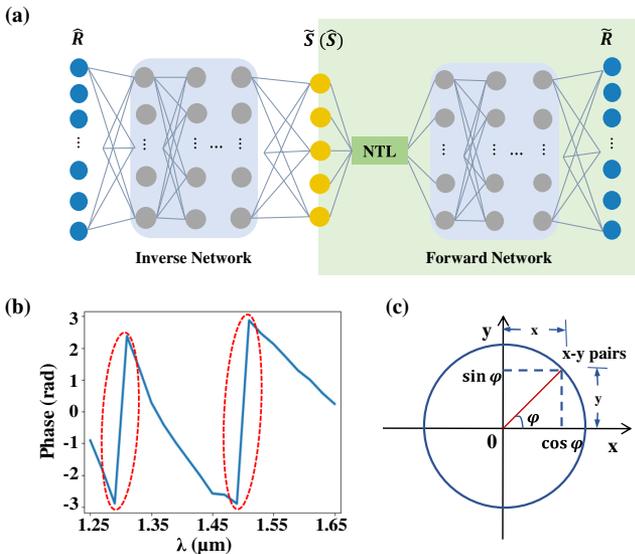
In our design, two kinds of meta-units, the elliptical nanopillar [2] and the nanofin [10] shown in Figs. 1(c) and 1(d), are simulated as datasets through the FDTD method. All meta-units are made of silicon, upon a quartz base. For the elliptical nanopillar, we sweep its major axis  $r_1$  and minor axis  $r_2$  from 0.2 to 0.8  $\mu\text{m}$  to get totally  $101 \times 101$  data pairs as dataset1. All elliptical nanopillars have a constant height of 1.4  $\mu\text{m}$  and a period of 0.9  $\mu\text{m}$ . For the nanofin, we fix its height and period to 1.4 and 0.85  $\mu\text{m}$ , respectively, and randomly pick its other five geometric parameters including widths ( $w_1, w_2$ ), lengths ( $l_1, l_2$ ), and the gap ( $g$ ). Totally, 24,000 data pairs of nanofins are simulated as dataset2.

The DNN is established with the tandem network [24,25] to achieve good convergence in the inverse design. As illustrated in Fig. 2(a), the DNN consists of a forward simulation network  $f_{\text{Forward}}$ , transforming input meta-unit parameters  $\hat{S}$  into corresponding EM responses  $\tilde{R}$ , and an inverse design network  $f_{\text{Inverse}}$ , aiming to predict meta-unit parameters  $\hat{S}$  according to required phases  $\hat{R}$ . The whole architecture of the DNN can be expressed as

$$Obj_{\text{Forward}}(w, b) = \text{argmin} \mathbf{E}(f_{\text{Forward}}(\hat{S}) - \hat{R})^2, \quad (1)$$

$$Obj_{\text{Inverse}}(w, b) = \text{argmin} \mathbf{E}(f_{\text{Inverse}}(\hat{R}) - \hat{S})^2, \quad (2)$$

where  $w$  and  $b$ , as the weight and bias, are parameters to be trained, and  $\mathbf{E}(\dots)^2$  is the mean square error (MSE) function.



**Fig. 2.** DNN architecture and characterization of phase information. (a) Architecture of the tandem DNN. The forward network transforms input meta-unit parameters  $\hat{S}$  into corresponding EM responses  $\tilde{R}$ . The inverse network predicts meta-unit parameters  $\hat{S}$  from required phases  $\hat{R}$ . (b) Example of the phase curve wrapped into the range  $[-\pi, \pi]$ . Phase jumps are circled with red dashed lines. (c) With the trigonometric function, the phase is transferred into the  $x - y$  projection pair.

In the forward network, a neural tensor layer (NTL) [26] is added to reconcile the huge dimension mismatch between meta-unit parameters and EM responses, which is especially beneficial to the training of dataset1. Fully connected layers compose other parts of the network.

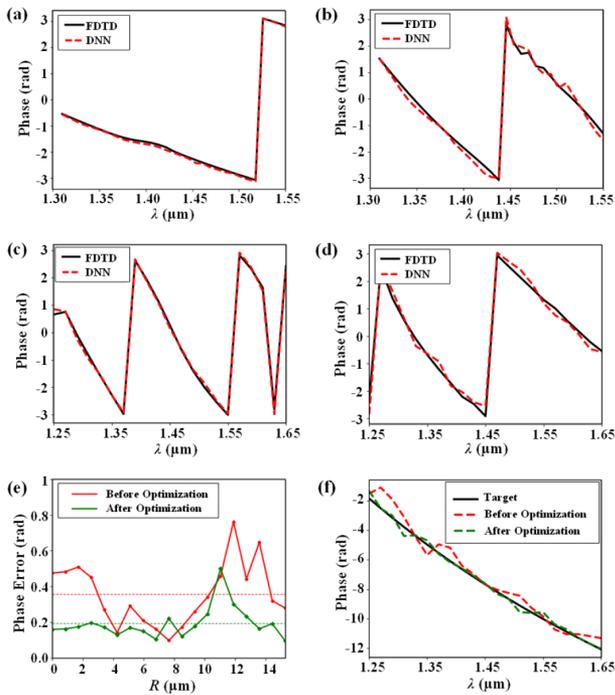
Figure 2(b) depicts the phase jump problem from the wrapped phase value gathered from the simulation. The abrupt phase change significantly disrupts the prediction of the DNN. To address this issue, past researchers trained and predicted the real and imaginary parts of the complex amplitude, respectively [27]. However, our target-fitting approach requires only the phase information and would even descend into phase chaos if fed with the transmittance coefficient. Here, illustrated in Fig. 2(c), we characterize the phase value by projecting the phase  $\varphi$  onto  $x$  and  $y$  axes through the trigonometric relation. The trigonometric transformation turns the jumping phase into the continuous  $x - y$  pair, which can be learned and predicted stably by the DNN. Meantime, the phase value can be exclusively retrieved from the  $x - y$  pair with the following formula, where  $x$  and  $y$  are the projection pair:

$$\varphi = \begin{cases} \arctan\left(\frac{y}{x}\right), & x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi, & x < 0, y \geq 0 \\ \arctan\left(\frac{y}{x}\right) - \pi, & x < 0, y < 0 \\ +\frac{\pi}{2}, & x = 0, y > 0 \\ -\frac{\pi}{2}, & x = 0, y < 0 \end{cases}. \quad (3)$$

During the training process, we accelerate the convergence with dynamic learning rates [28] and improve the DNN performance with batch normalization [29]. Weights of the network are adjusted incrementally through the Adam algorithm. After the forward network is well trained with data pairs  $(\hat{S}, \hat{R})$ , the weights are fixed. Then the inverse network is trained to minimize the difference between  $\hat{R}$  and  $\tilde{R}$ . The meta-unit parameter  $\hat{S}$  is generated from the intermediate layer of the whole network. The performance of the DNN is evaluated on the testing dataset, through the loss function defined as the MSE between the ground truth (simulated  $x - y$  pairs) and the predicted output (predicted  $x - y$  pairs). After 500 epochs,  $x - y$  MSE of the tandem network is 0.098 on dataset1 and 0.147 on dataset2. Some examples of the trained phase curves transferred from the  $x - y$  pairs are showed in Figs. 3(a)–3(d). Phases obtained from the DNN agree well with those simulated by the FDTD method.

Here, besides those mentioned in previous researches, we further discuss other advantages of the tandem network. The tandem architecture performs better than a single inverse network because it not only promotes the convergence of training, but also captures the essence of inverse design problems. When learning the relationship between EM responses and meta-units with only a single inverse network, even though a low loss is achieved, the distinction between predicted and target EM responses remains obvious. This is because minute errors of geometric parameters of meta-units can lead to overall response shifts. On the contrary, the tandem network avoids such intermediate errors by directly building the relationship between demanding responses and real responses of predicted meta-units.

Considering the small amount of targeting features, a two-step optimized training process containing a fine-tuning [30] procedure is applied to improve the precision. We utilize the



**Fig. 3.** Several results from the trained and optimized DNN. (a), (b) Examples of inverse network training results based on dataset1. (c), (d) Examples of inverse network training results based on dataset2. In (a)–(d), red dashed lines are phase curves predicted by the DNN, and black solid lines are simulation results. (e) Phase errors before (red line) and after (green line) DNN optimization in different radial positions. Each point represents sum of phase error of all wavelengths on a certain radial position. After optimization, the average phase error of one wavelength is 0.0093 rad. (f) Example of the target fitting method. Black solid line is the target phase, while red and green dotted lines are, respectively, the responses of the meta-units generated from the DNN before and after optimization.

trained network as a pre-trained model, then fine-tune the inverse network in a narrow data space where the data are similar to target features. The fine-tuning process can be explained as modifying the encoder (inverse network) with an ideal decoder (well-trained and fixed forward network) in a specific dataspace. The decline of the loss value in the narrow data space represents the further optimization of the inverse network. We manually generate phase curves as the dataset for fine-tuning, taking advantage of the tandem architecture with which the input EM response can be transferred into the output phase without meta-unit parameters. These manual phase curves are empirically generated by adjusting interceptions of target phase curves.

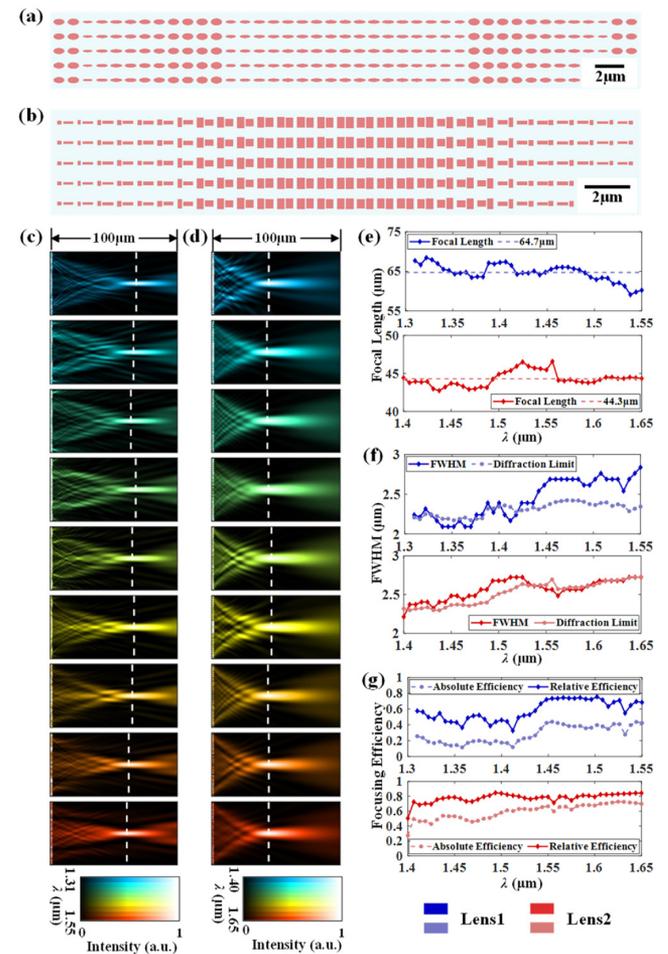
After the whole DNN is thoroughly trained and optimized, the target-fitting method can guide the achromatic design following the phase profile

$$\Delta\varphi(r, \lambda) = -\frac{2\pi}{\lambda} \left( \sqrt{r^2 + f^2} - \sqrt{r_0^2 + f^2} \right) + C_0, \quad (4)$$

where  $r$  is the radial position of the metalens,  $\lambda$  is the operating wavelength, and  $f$  is the focal length.  $r_0$  and  $C_0$  are two constants determining the initial phase at  $r = 0$ . We target two 1D achromatic metalenses, separately based on elliptical nanopillars (lens1) and nanofins (lens2). The operating bandwidth, focal lengths, and radii of two metalenses are set as 1.31 ~ 1.55  $\mu\text{m}$

and 1.40 ~ 1.65  $\mu\text{m}$ ,  $f = 70 \mu\text{m}$  and  $f = 50 \mu\text{m}$ , and  $r = 18.5 \mu\text{m}$  and  $r = 12.3 \mu\text{m}$ . Corresponding numerical apertures are NA = 0.26 and NA = 0.24 for lens1 and lens2, respectively.  $r_0$  and  $C_0$  are optimized to be 36.90 and 1.26 for lens1, and 30.18 and -0.08 for lens2. The geometric parameters of all meta-units are inversely generated through the DNN according to the required wavelength dependent phases in every position. Figure 3(e) summarizes phase errors of lens2 in different radial positions before and after optimization of the DNN. Compared with the pre-trained network, the optimized network obtains more accurate inverse design results. A fitting example before and after optimization is shown in Fig. 3(f), indicating an obvious improvement of the overall fitting performance through optimization.

To further verify our design, these two 1D metalenses are simulated by the FDTD method. Figures 4(a) and 4(b) are the 1D nanostructure arrangements of two lenses. Figures 4(c) and 4(d) depict simulated normalized intensity profiles in the  $x - z$  plane at nine selected wavelengths. Focal lengths of lens1 and lens2 both exhibit weak dependence on the wavelength. As



**Fig. 4.** 1D nanostructure arrangements and simulation results of achromatic metalenses made of elliptical nanopillars (lens1) and nanofins (lens2). (a) 1D nanostructure arrangement of lens1. (b) 1D nanostructure arrangement of lens2. (c) Normalized intensity profiles in the  $x - z$  plane of lens1 at nine wavelengths, within the operating bandwidth from 1.31 to 1.55  $\mu\text{m}$ . White dashed lines are focal planes. (d) Normalized intensity profiles of lens2 at nine wavelengths, within the operating bandwidth from 1.4 to 1.65  $\mu\text{m}$ . (e) Focal lengths, (f) FWHMs, and (g) focusing efficiencies of two metalenses.

