Inverse design of ultracompact multi-focal optical devices by diffractive neural networks

Yuyao Chen, Yilin Zhu, Wesley A. Britton, and Luca Dal Negro

1Department of Electrical and Computer Engineering and Photonics Center, Boston University, 8 Saint Mary’s Street, Boston, MA, 02215, USA
2Division of Materials Science and Engineering, Boston University, 15 St Mary’s St, Brookline, MA, 02246, USA
3Department of Physics, Boston University, 590 Commonwealth Avenue, Boston, MA, 02215, USA

We propose an efficient inverse design approach for multifunctional optical elements based on adaptive deep diffractive neural networks (a-D^2NNs). Specifically, we introduce a-D^2NNs and design two-layer diffractive devices that can selectively focus incident radiation over two well-separated spectral bands at desired distances. We investigate focusing efficiencies at two wavelengths and achieve targeted spectral lineshapes and spatial point-spread functions (PSFs) with optimal focusing efficiency. In particular, we demonstrate control of the spectral bandwidths at separate focal positions beyond the theoretical limit of single-lens devices with the same aperture size. Finally, we demonstrate devices that produce super-oscillatory focal spots at desired wavelengths. The proposed method is compatible with current diffractive optics and doublet metasurface technology for ultracompact multispectral imaging and lensless microscopy applications.

Multifunctional diffractive optical elements (DOEs), when integrated atop on-chip detectors, enable ultracompact imaging functionalities for miniaturized flat cameras and microscopes [1-4]. Multispectral behavior is often achieved by partitioning single layer devices into separate phase regions that affect different wavelengths. However, this design limits the maximum efficiency achievable at each wavelength, which is a significant challenge for DOEs working at multiple wavelengths [5, 6]. This is because when one specific wavelength illuminates the entire device, only the phase region designed to operate at that wavelength will produce the desired output while the other part of the illuminated device area will not, thus requiring a different approach.

In order to address this important challenge, we propose here novel multi-layer designs based on the flexibility of adaptive diffractive neural networks (a-D^2NNs) for the engineering of multi-layered diffractive devices with targeted spectral response and spatial point-spread functions (PSFs) at different wavelengths. Recently, deep diffractive neural networks (D^2NNs) that combine optical diffraction with deep learning capabilities have been reported and applied to all-optical diffraction-based systems that implement object recognition [7]. Moreover, D^2NNs have also been demonstrated successfully for the inverse design of multi-layered diffractive devices that achieve pulse shaping [8] and broadband filtering [9]. These devices are macroscopic with typical dimensions up to the cm size and are fabricated using 3D printing for applications in the Terahertz domain [8, 9]. However, the design of diffractive devices that operate over multiple spectral bands in the optical regime is very challenging and requires a more flexible implementation of the D^2NNs platform.

In this paper, we introduce and utilize a-D^2NNs that leverage an adaptive loss weight algorithm for the inverse design of two-layer, ultracompact dual-band DOEs. The a-D^2NNs are trained to maximize the focusing efficiencies for \( \lambda_1 \) at \( f_1 \) and \( \lambda_2 \) at \( f_2 \). The engineered devices show efficiencies over 50% at both targeted wavelengths, which exceeds the limit of phase-modulated single layer DOEs [5, 6, 10]. We systematically investigate how the focusing efficiencies vary with the distance between the two diffractive layers and the pixel size, taking into account practical fabrication constraints. We also investigate how the efficiency is affected by the phase discretization level of the proposed diffractive devices.

The obtained phase designs can also be implemented using current metasurface technology [1, 11-13], including the recently developed doublet metasurface fabrication approach [14, 15]. An important aspect of our approach is the design of the spectral lineshapes of DOEs. In fact, we demonstrate dual-band devices with designed bandwidths that are narrower compared to diffractive lenses with the same aperture size. Finally, we show that a-D^2NNs can be implemented to design devices with desired spatial PSFs, including DOEs that produce superoscillatory fields with focal spots below the diffraction limit [16].

Figure 1 (a) illustrates the general two-layer diffractive device concept consisting of two diffractive phase plates located on both sides of a transparent substrate. The varying thickness profiles of the materials on the phase plates impart different phase shifts to the waves that propagate through the device. A schematic design of the a-D^2NN that implements such a device is shown in Fig. 1 (b), where the two diffractive layers of the a-D^2NN correspond to the two phase plates of the device. We implement the Rayleigh-Sommerfeld (RS) first integral formulation within the a-D^2NN in order to simulate the forward light propagation from one plane to the next...
and the index of medium between the two planes. We use where \( \lambda \) is the incident wavelength in vacuum and \( n \) is the index of medium between the two planes. In our two-layer DOE, we first compute the field distributions on the source and observation plane with coordinates \((x, y)\) and \((x', y')\), respectively. Moreover, \( k = \frac{2\pi}{\lambda} \) is the wave number, where \( \lambda \) is the incident wavelength in vacuum and \( n \) is the index of medium between the two planes. We use \( r = \sqrt{x'^2 + y'^2 + z^2} \), where \( z \) is the distance between the two planes. In our two-layer DOE, we first compute the forward propagation from plane \( z = 0 \) to plane \( z = d \) at wavelengths \( \lambda_1 \) and \( \lambda_2 \). Then the field distributions on the focal plane for \( \lambda_1 \) at \( z = d + f_1 \) and for \( \lambda_2 \) at \( z = d + f_2 \) are calculated. We then utilize the a-D\(^2\)NN to maximize the focusing efficiency \( \eta \) at these two focal planes, using the following definition for the focusing efficiency [10]:

\[
\eta(\lambda, z) = \frac{\int_0^{2\pi} \int_0^{\text{FWHM}/2} I'(\lambda, z, \rho', \theta') dr'd\theta'}{\int \int I(\lambda, z = 0, \rho, \theta) dS}
\]

where \( I' \) denotes the intensity distribution on the focal plane, \( I \) denotes the one on the input plane, and \( S \) denotes the input plane aperture. The symbols \((\rho, \theta)\) and \((\rho', \theta')\) are the polar coordinates on the focal and input plane, respectively.

The focusing efficiency is utilized in the loss function of the a-D\(^2\)NN as follows:

\[
\mathcal{L} = w_1(1 - \eta_1)^2 + w_2(1 - \eta_2)^2
\]

where \( \eta_1 = \eta(\lambda_1, d + f_1) \), \( \eta_2 = \eta(\lambda_2, d + f_2) \), and \( w_1 \) and \( w_2 \) are the loss weights. Based on the definition of a suitable loss function, the a-D\(^2\)NN is directly trained using error backpropagation within the diffractive layers without the need of training datasets. Therefore, the a-D\(^2\)NN achieves a more efficient inverse design of complex phase devices compared to data-driven neural network approaches [17, 18]. Specifically, a-D\(^2\)NN are trained by varying the phase profiles on the two diffractive layers in order to minimize \( \mathcal{L} \). We train the latent variable \( h_{\ell} \) on each pixel of the diffractive layers related to the material thickness \( h \) by \( h = h_{\text{max}}(\sin(h_{\ell}) + 1)/2 \), where \( h_{\text{max}} \) is the specified maximum thickness of the device. The phase profile \( \phi(x, y) \) induced by diffractive layers at wavelength \( \lambda = \frac{2\pi}{\lambda_{\text{eff}}}(n - 1)h \). As a proof of concept, we select \( \lambda_1 = 632.8 \text{ nm} \), \( \lambda_2 = 808 \text{ nm} \), \( f_1 = 80 \mu\text{m} \), and \( f_2 = 110 \mu\text{m} \), which were used in our previous work [10]. The two diffractive layers are square apertures with \( L = 100 \mu\text{m} \) side length and the device pixel size is \( \Delta x = 200 \text{ nm} \). We assume that the substrate index \( n = 3 \), \( h_{\text{max}} = 500 \text{ nm} \), and thickness \( d = 200 \mu\text{m} \). Differently from the usual D\(^2\)NN approach, here we implement adaptive loss weights that balance the interplay between different loss terms automatically depending on their values [19, 20]. In particular, we apply the following updates for \( w_{\ell}(m = 1, 2) \) at the \( k^{th} \) epoch of training:

\[
w_{\ell}^k ← w_{\ell}^{k-1} + \gamma(1 - \eta_{m})^2
\]

where \( \gamma \) is the learning rate for loss weights update and we choose \( \gamma = 1 \). We use a desktop with GeForce GTX 1080 Ti graphical processing unit (GPU, Nvidia Inc.), an Intel i7-8700K central processing unit (CPU, Intel Inc.) and 32 GB of RAM for training. We trained the a-D\(^2\)NN over 2000 epochs using Adam optimizer with learning rate equals to 0.1. The typical training time is only \( \sim 10 \) minutes. In Fig. 1 (c) we show the obtained focusing efficiency spectra of the device at \( f_1 \) and \( f_2 \). Specifically, we observe that \( \eta(f_1, \lambda) \) and \( \eta(f_2, \lambda) \) are peaked at \( \lambda_1 \) and \( \lambda_2 \), respectively, and that both \( \eta_1 \) and \( \eta_2 \) values exceed 50%. Therefore, the designed two-layer dual-band device exceeds the efficiency limit expected in a single-layer DOE. Moreover, in Fig. 1 (d) we display the side view of the normalized intensity diffraction of the device,

FIG. 1. (a) Schematics of a two-layer dual-band DOE and (b) D\(^2\)NN representation. (c) Focusing efficiency spectra for the device in panel (a) with a-D\(^2\)NN. (d) Side view of normalized diffraction intensity at \( \lambda_1 \) and \( \lambda_2 \).

FIG. 2. (a) Focusing efficiency with respect to the distance \( d \) between the two diffractive layers (b) Focusing efficiency with respect to the number of discrete phase levels. The inset shows the dependence of the focusing efficiencies on the pixel size (minimum spatial feature of the phase profile).
and $\eta$ between the two diffractive layers while keeping all the $f_1$ and $\lambda_2$ respectively.

We further investigate the influence of the distance $d$ between the two diffractive layers while keeping all the other parameters as specified above. The obtained $\eta_1$ and $\eta_2$ for devices with different $d$ are shown in Fig. 2 (a), which demonstrates that for all the devices the values of $\eta_1$ and $\eta_2$ remain above 50%. We next consider the discretization of the obtained continuous phase profiles into discrete levels used for scalable lithographic fabrication [2, 10]. Figure. 2 (b) displays how the number of discretized levels of the device affects $\eta_1$ and $\eta_2$. In particular, we find that $\eta_1$ and $\eta_2$ increase with the number of phase levels used, with a similar scaling to the one reported for the diffraction efficiency of multi-level gratings [21]. In particular, we observe that our two-layer device can achieve $\eta_1 = 46\%$ and $\eta_2 = 52\%$ when only 8 discretization levels are used. The inset of Fig. 2 (b) displays the focusing efficiency with respect to the pixel size $\Delta x$. Devices with smaller $\Delta x$ achieve larger $\eta_1$ and $\eta_2$ efficiencies as they accommodate faster phase variations. Therefore, our analysis indicates that the proposed devices can be conveniently engineered using current multi-level DOE technology [2, 10] as well as planar metasurfaces that provide nanoscale phase resolution (50 nm ~ 300 nm) [11, 12].

Another important advantage of the DOE design based on a-D$^2$NNs is that we can engineer spectral lineshapes by modifying the loss function used for training the network. In order to demonstrate this capability, we train a-D$^2$NNs to obtain devices with spectral lineshapes for the focusing efficiency $\eta_m(f_m, \lambda)$ $(m = 1, 2)$ described by the expression:

$$\eta_m(f_m, \lambda) = \exp \left[ -4\log(2) \left( \frac{|\lambda - \lambda_m|}{\sigma_m} \right)^2 \right] \quad (m = 1, 2)$$

where $\lambda_m$ and $\sigma_m$ $(m = 1, 2)$ are the center wavelength and FWHM of the targeted Gaussian spectral lineshape, respectively. We modify the loss function in a-D$^2$NN as follows:

$$\mathcal{L} = \sum_{m=1,2} \sum_{k=1}^{N} w_m(1 - \eta_m)^2 + \frac{1}{N} w_{sm} [\eta_m*\eta_m(f_m, \lambda_k^m) - \eta_m(f_m, \lambda_{k}^m)]^2$$

where $w_{sm}$ $(m = 1, 2)$ is the loss weight for the spectral lineshape loss term. The first term is the same used in Eq. 4. For the second term, we sample $N$ discrete wavelengths of the target spectrum uniformly from $\lambda_{m, \min}$ to $\lambda_{m, \max}$ centered at $\lambda_m$ $(m = 1, 2)$ and evaluate the focusing efficiencies over these wavelengths. The mean squared error (MSE) between the obtained $\eta_m(f_m, \lambda_k^m)$ and the target lineshape $\eta_m(f_m, \lambda)$ with its maximum rescaled to $\eta_m$ is then calculated. During the training process, we apply the adaptative loss weights for both $w_m$ and $w_{sm}$. We train the a-D$^2$NN with the same parameters used to generate Fig. 1 and we fix $\sigma_1 = \sigma_2 = \sigma$. In particular, we sampled $N = 30$ wavelengths over two ranges between $\lambda_{m, \min} = \lambda_m - 3\sigma$ and $\lambda_{m, \max} = \lambda_m + 3\sigma$. The a-D$^2$NN is trained over 2000 epochs. We show the spectral results for the device trained using $\sigma = 10$ nm and $\sigma = 40$ nm in Fig. 3 (a) and (b), respectively. Furthermore, in Fig. 3 (c) we display how $\eta_1$ and $\eta_2$ vary for devices optimized with different $\sigma$. A sharp drop of focusing efficiency is observed when the width of the targeted Gaussian lineshape is decreasing below 20 nm. We also evaluate the normalized field intensity spectra at the origin of focal planes at $f_1$ and $f_2$ for the device with $\sigma = 10$ nm. We compare our results with the ones of two diffractive lenses with the same dimension that focus $\lambda_1$ at $f_1$ and $\lambda_2$ at $f_2$. The analytical expression for the normalized intensity spectrum $I_m(\lambda)$ of a diffractive lens that focuses $\lambda_m$ at $f_m$ $(m=1,2)$ is given by [22]:

$$I_m(\lambda) = \left[ \frac{\sin(u_m(\lambda)/4)}{u_m(\lambda)/4} \right]^2$$

where we defined $u_m = \frac{2\pi c}{\lambda} \left( \frac{\lambda}{f_m \lambda_m} - \frac{1}{f_m} \right)$. As shown in Fig. 3 (d), two-layer devices designed using the a-D$^2$NN method can feature significantly narrow bandwidths than diffractive lenses at both $\lambda_1$ and $\lambda_2$. Recalling that the intensity spectrum for diffractive lens is completely determined once the parameters $L$, $\lambda_m$, and $f_m$ are fixed, we appreciate that the designed two-layer DOEs provide the additional capability to tailor spectral lineshapes for a given aperture size. To better understand how the obtained devices achieve narrow band-
FIG. 4. (a) Normalized transverse intensities at \(\lambda_1\) (red solid line) and (b) at \(\lambda_1\) (blue solid line) and corresponding phase gradients across the focal spots. (c, d) Focal plane intensity profiles at \(\lambda_1\) and \(\lambda_2\), respectively, with \(\alpha = 0.4\). (e, f) Normalized phase gradient maps at the focal planes for \(\lambda_1\) and \(\lambda_2\), respectively. For better visualization, the range of the color bars is limited to \([0, 1]\). (g, h) \(\eta_1\) and \(\eta_2\) with respect to \(\alpha\).

widths, we evaluate the spectral dependence of their focal lengths near \(\lambda_1\) and \(\lambda_2\) and compare with diffractive lenses in Fig. 3 (e) and (f), respectively. Our findings show that the focal lengths of the designed dual-band devices vary faster with respect to the wavelength compared to diffractive lenses. Therefore, they can achieve enhanced spectral selectivity (narrower bandwidths) due to their stronger defocusing behavior when varying the incident wavelengths.

We finally implement a-D^2NNs for the inverse design of two-layer DOEs with desired focusing PSF. We model the PSF by two-dimensional Gaussian function on the focal plane \(f_m\) (\(m = 1, 2\)):

\[
I_{tm}(x, y, z_m, \lambda_m) = \exp \left[-4 \log(2) \left(\frac{x^2 + y^2}{(\alpha \eta_m)^2}\right)\right]
\]

where \(z_m = d + f_m\) are the focal plane \(z\) coordinate for \(\lambda_m\), \(\alpha\) is a scaling constant that quantifies the degree of spatial localization of the designed focal spot with respect to the Rayleigh diffraction limit, which is achieved for \(\alpha = 1\), and \(\eta_m = 0.51 \lambda_m f_m\) is the diffraction limited FWHM of the focal spot. In order to obtain desired spatial PSFs we implement the following loss function for training:

\[
\mathcal{L} = \sum_{m=1,2} \sum_{x,y} w_m (1 - \eta_m)^2 + w_{pm} \left[I_m(0, 0, z_m, \lambda_m) \times I_{tm}(x, y, z_m, \lambda_m) - I_m(x, y, z_m, \lambda_m)\right]^2
\]

where \(w_{pm}\) is the loss weight for the loss term of squared error between device real PSF \(I_m(x, y, z_m, \lambda_m)\) and targeted PSF \(I_{tm}(x, y, z_m, \lambda_m)\) with its maximum rescaled to \(I_m(0, 0, z_m, \lambda_m)\). In particular, we trained a-D^2NN using \(\alpha = 0.4\), which corresponds to a FWHM below the diffraction limit. The intensity cuts through the center of the focal spots at \(\lambda_1\) and \(\lambda_2\) are shown in Fig. 4 (a) and (b), respectively. The dashed lines are the targeted PSF used for training the network. The obtained intensity profiles indicate the formation of optical super-oscillations, which have been shown to result in arbitrarily small energy concentration without the assistance of evanescent waves [23, 24]. We note that the obtained PSFs exhibit the presence of significant sidebands compared to the targeted Gaussian PSF. This is due to the fundamental nature of super-oscillations in which enhanced (sub-diffractive) field focusing can only be achieved at the expense of a polynomial increase in the power directed into the sidebands [16]. Due to their extreme localization properties, optical super-oscillations have found applications to sub-wavelength imaging and microscopy [25]. In order to demonstrate super-oscillations in our devices we studied the phase gradient \(|\nabla \phi|\) of the diffracted field on the focal plane, which corresponds to a local wave number. Super-oscillations form when \(|\nabla \phi| > k_m\), where \(k_m = 2\pi \lambda_m/(m = 1, 2)\) are the incident wave numbers. In Fig. 4 (a) and (b) we display for the two wavelengths of interest the phase gradient profiles of the fields normalized by \(k_m\). We notice that the peaks of \(|\nabla \phi|/k_m\) exceed unity around the designed focal spots, demonstrating the super-oscillation character of the waves. We further show the two-dimensional focal intensity distributions and \(|\nabla \phi|/k_m\) maps on the two different focal planes in Fig. 4 (c-d) and (e-f), respectively. In Fig. 4 (g) and (h) we summarize our results for the variation of focusing efficiencies with respect to the localization parameter \(\alpha\). We note that the focusing efficiencies slightly decrease when increasing \(\alpha\) if \(\alpha > 1\) while they suddenly drop to almost zero by decreasing.
When $\alpha < 1$, consistently with the super-oscillating regime [23, 24].

In conclusion, we introduced an inverse design approach for dual-band multi-focus DOE based on flexible a-D2NNs. We demonstrate novel two-layer designs that show $\eta_1, \eta_2 > 50\%$, beyond the limit of single-layer DOEs working at two wavelengths. Furthermore, we showed the designs of DOEs with desired spectral lineshapes and FWHM down to $\sigma = 5 \text{ nm}$. Finally, we show PSF engineering with designed super-oscillatory focal spots below the diffraction limit. The flexible approach introduced here enables the engineering of two-layer diffractive devices with desired spectral and spatial responses for multi-band imaging and microscopy applications.

FUNDING INFORMATION

National Science Foundation (ECCS-2015700).

DISCLOSURES

The authors declare no conflicts of interest.
on signal processing 54, 3732–3740 (2006).

[25] Edward TF Rogers, Jari Lindberg, Tapashree Roy, Salvatore Savo, John E Chad, Mark R Dennis, and Nikolay I Zheludev, “A super-oscillatory lens optical microscope for subwavelength imaging,” Nature materials 11, 432–435 (2012).