ARTICLE

Check for updates

Large area optimization of meta-lens via data-free machine learning

Maksym Zhelyeznyakov [1]✉, Johannes Fröch[1,2], Anna Wirth-Singh[2], Jaebum Noh[3], Junsuk Rho [3,4,5], Steve Brunton [6] & Arka Majumdar [1,2]✉

Sub-wavelength diffractive optics, commonly known as meta-optics, present a complex numerical simulation challenge, due to their multi-scale nature. The behavior of constituent sub-wavelength scatterers, or meta-atoms, needs to be modeled by full-wave electromagnetic simulations, whereas the whole meta-optical system can be modeled using ray/Fourier optics. Most simulation techniques for large-scale meta-optics rely on the local phase approximation (LPA), where the coupling between dissimilar meta-atoms is neglected. Here we introduce a physics-informed neural network, coupled with the overlapping boundary method, which can efficiently model the meta-optics while still incorporating all of the coupling between meta-atoms. We demonstrate the efficacy of our technique by designing 1mm aperture cylindrical meta-lenses exhibiting higher efficiency than the ones designed under LPA. We experimentally validated the maximum intensity improvement (up to 53%) of the inverse-designed meta-lens. Our reported method can design large aperture ($\sim 10^4 - 10^5 \lambda$) meta-optics in a reasonable time (approximately 15 minutes on a graphics processing unit) without relying on the LPA.

[1] Department of Electrical and Computer Engineering, University of Washington, Seattle 98195 Washington, USA. [2] Department of Physics, University of Washington, Seattle 98195 WA, USA. [3] Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. [4] Department of Chemical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. [5] POSCO-POSTECH-RIST Convergence Research Center for Flat Optics and Metaphotonics, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. [6] Department of Mechanical Engineering, University of Washington, Seattle 98195 WA, USA. ✉email: mzhelyez@gmail.com; arka@uw.edu

In the age of silicon computing, numerical simulations are at the heart of understanding and designing physical systems. For many cases, analytical solutions to complex device geometries are intractable to compute, or simply do not exist. From extremely large systems like rockets[1] to ultra-small nanophotonic devices[2], numerical simulations provide scientists and engineers with the necessary tools to design nonintuitive structures. In electromagnetics, direct solvers, including the finite difference time domain (FDTD)[3] and the finite difference frequency domain (FDFD)[4,5] simulators, are the usual choices when dealing with heterogeneous structures with subwavelength features that require a high degree of numerical accuracy. Most commonly, electromagnetic simulation tools serve to validate the qualitative designs created by engineers based on prior knowledge and intuition. In recent years, the field of nanophotonics has incorporated a new paradigm of computer-aided device design, where a device's performance is summarized by a quantitative figure of merit (FOM) that is optimized over. This method involves running a forward numerical simulation, computing the FOM, and iteratively modifying the device's geometry based on an optimization algorithm to reach the desired FOM. Such optimization methods, often termed as inverse design, have already been used to create multi-functional and efficient nanophotonic structures[2,6-15]. However, electromagnetic simulators suffer from a computational resource problem when the device dimension becomes large ($\gtrsim 10^3 \lambda$), where λ is the device's operating wavelength. As most electromagnetic simulations are performed over a sub-wavelength grid size, with increased size, the number of input variable becomes prohibitively large, making the simulation slow and memory extensive. The limitation of such forward electromagnetic simulators becomes even more severe for inverse design, where many such forward simulations are needed.

Sub-wavelength diffractive optics, also known as meta-optics, present an important test-bed for these problems: the constituent elements of the meta-optics, i.e. meta-atoms, are sub-wavelength, but the dimensions of the whole meta-optics are on the order of $\sim 10^3 \lambda - 10^5 \lambda$. Thus the underlying physics of each scatterer has to be modelled using full-wave electromagnetic simulation, but the whole meta-optical system needs to be simulated using ray or wave optics. Such multiscale electromagnetic simulators invariably rely on approximations, the most common of which is the local phase approximation (LPA): the scattering in any small region is taken to be the same as the scattering from a periodic surface[9]. This approximation allows the simulation of each scatterer in a periodic array, abstracting out the electromagnetic response as a simple phase shift. While this significantly reduces the computational complexity of simulating a meta-optic, this approximation fails to consider the coupling of each scatterer with their dissimilar neighbors. In fact, it has already been shown that meta-optical lenses designed under LPA have suboptimal efficiencies[16], especially when the numerical aperture is large. The LPA becomes even more inaccurate when the material used to create the meta-lens has low index, such as SiN[17]. We note that, while a full FDTD coupled with adjoint optimization has been used to design a meta-optic without relying on LPA, their size has been limited to only ~100λ[18]. LPA can also be bypassed using Mie scattering approaches[19], which however limits the shape of scatterers.

To address the computational bottleneck of large-area inverse design, here we introduce a physics-informed neural network (PINN), to model super-cell subsections of a larger metasurface[20-22] which in conjunction with the overlapping boundary method[23-27], can replace a traditional FDTD/ FDFD solver to predict the electric field distribution for a given dielectric distribution. PINNs and other model-based deep learning architectures have already been used in modeling physical systems[28].

We also note that a large number of works already used artificial neural networks to predict spectral responses of meta-optics of varying scatterer geometries[25,29-36]. However, these works used largely periodic structures for which LPA is accurate. We present a solution via PINNs[37,38] for lenses and devices with spatially varying scatterer geometries, where it is necessary to model the whole electric field from several scatterers and their neighbors. The use of PINNs to accurately model the electromagnetic scattering beyond the LPA is the main contribution of this work. PINNs solve partial differential equations (PDEs) by minimizing a loss function constructed from the PDE itself. This loss function is generally some norm of the residual[37] or an energy function derived from the PDE[39]. PINNs have already seen wide usage in the field of fluid mechanics[40-42], biology[43], and solving stochastic PDEs[44]. In electromagnetic inverse problems, PINNs have also been employed to design meta-optics and nanophotonic devices[45,46]. These works, however, did not clearly demonstrate a simulation speedup, and are limited to the inverse design of only very small devices. We also note that pre-trained PINNs have been used to design small gratings[47]; however their methodology is limited to small gratings that deflect light fields to specific angles, and thus cannot be readily used for the inverse design of arbitrary meta-optics or a meta-lens.

In our work, we train PINNs to predict the electric fields from a parameterized set of dielectric meta-atoms corresponding to rectangular pillars. We then use this as a surrogate model to design cylindrical meta-lenses operating in the visible with a diameter of 1 mm (~1500λ). Large area meta-optics are simulated by partitioning the simulation region into groups of 11 meta-atoms, with the outermost meta-atoms overlapping. After simulation, the fields are stitched together. Our PINNs do not require a training data set. They are trained by randomly generating distributions of dielectric meta-atoms ϵ, feeding them into a neural network NN, and minimizing the residual of the linear Maxwell PDE operator

$$\left\| A_{\text{Maxwell}}(\epsilon)NN(\epsilon) - b \right\|_1 \tag{1}$$

over the neural network training parameters. This means our PINNs are trained without ever invoking a forward numerical simulation of Maxwell's equations during the training process. Numerical simulations are invoked only to test the neural network performance (see next section, Supplementary Note 6, and Supplementary Fig. 5). A similar data-free approach has been applied to deep-tissue microscopy[48], however inverse design was not demonstrated. Once trained, this method can calculate the full electromagnetic field response from a 1 mm diameter cylindrical meta-lens at ~630nm in approximately 3 seconds on a graphics processing unit (GPU). Furthermore, we demonstrate an experimental improvement (over 50%) of the maximum intensity of cylindrical metalenses over their forward designed hyperboloid counterparts, signifying the improvement over using LPA. We note that the reported method is robust enough to handle even larger meta-optics, with simulation time scaling only linearly with the aperture of the cylindrical lens (see Supplementary Note 9).

Methods

Deep neural network proxy to Maxwell's equations. Our problem statement is summarized in Fig. 1c. The monochromatic electromagnetic scattering equation for an inhomogeneous, nonmagnetic material is given by:

$$\nabla \times \nabla \times \mathcal{E}(x) - \omega^2 \epsilon(x)\mathcal{E}(x) = i\omega \mathcal{J}(x). \tag{2}$$

In the 2D case, assuming out of plane polarization $(0, 0, \mathcal{E}_z)$, and the double curl vector identity, $\nabla \times \nabla \times = \nabla (\nabla \cdot) - \nabla^2$ we can
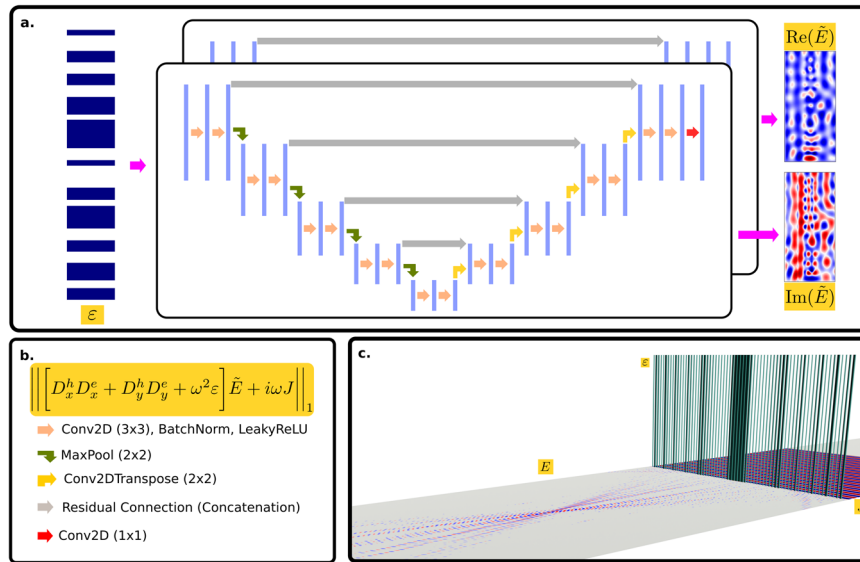
Fig. 1 Problem outline. a Neural network schematic. (For a more detailed schematic see Supplementary Fig. 10). ε distributions of 11 pillar meta-optics are meshed by randomly generating sets of pillar half-widths of height $h = 0.6\mu m$ with a dielectric constant 4 corresponding to SiN. The background medium is air. The loss function is the $|| \cdot ||_1$ norm of the residual of Eq. (2). **b** Neural network architecture. Encoder layers are down-sampled by a maxpool operation with a 2×2 kernel. The decoder part of the network is up-sampled by the Conv2DTranspose operation with a 2×2 kernel. **c** Render of the system under optimization. A current J is incident on a cylindrical metalens with dielectric distribution ε, with output response E.

simplify Eq. (2) to:

$$\nabla^2 \mathcal{E}_z(x) + \omega^2 \epsilon(x) \mathcal{E}_z(x) = -i\omega \mathcal{J}_z \quad (3)$$

where $\mathcal{E}_z$ and $\mathcal{J}_z$ are scalar fields. Equation (3) is defined over all space, with boundary conditions at $|x| \to \infty$. To simulate this equation, we discretize it on a Yee grid[3] by replacing the ∇ operator with a matrix, and treating the field $\mathcal{E}_z(x)$ and current $\mathcal{J}_z$ as vectors E and J at discrete values of x. Similarly, we treat the dielectric distribution $\epsilon(x)$ as a diagonal matrix ε. To truncate the simulation to a finite domain, we use perfectly matched boundary layers (PML), by making the transformation on the partial derivative operators $\frac{\partial}{\partial x} \to \frac{1}{1+i\frac{\sigma(x)}{\omega}}\frac{\partial}{\partial x}$. Making these substitutions, Eq. (3) becomes:

$$\left[D_x^h D_x^e + D_y^h D_y^e + \omega^2 \varepsilon \right] E = -i\omega J \quad (4)$$

with matrices $D_x^h, D_x^e, D_y^h, D_y^e$ being the matrix representations of corresponding derivative operators on a Yee grid with incorporated PML boundaries. See Supplementary Note 5 and Supplementary Fig. 4 for a more detailed description of the matrices. These matrices were extracted from a modified version of the package angler[49] with constants c, ϵ, μ set to 1 and the length scale set to μm. To build a neural network proxy to solve Eq. (4), we employ a PINN (Fig. 1a and b). PINNs generally use the coordinates of the computational grid as the input to the neural network, and then minimize the residual of the physical equations by approximating the target quantity being solved for with a neural network. This approach is slow since it effectively functions as an iterative solver re-parameterized over neural network weights and biases. It also required retraining the neural network for all different dielectric distributions. Our approach is to build a proxy solver that predicts the field E from a dielectric distribution ε. We pretrain the PINN to predict fields from inputs ε before optimizing our meta-lenses. The minimization problem to train

the PINN becomes:

$$\min_\theta f(\varepsilon; \theta)$$

$$\text{where} \quad f(\varepsilon; \theta) = \left\| \left[D_x^h D_x^e + D_y^h D_y^e + \omega^2 \varepsilon \right] NN(\varepsilon; \theta) + i\omega J \right\|_1 \quad (5)$$

with $NN(\varepsilon; \theta)$ being the output field from the PINN, and $|| \cdot ||_1$ is the vector l_1 norm. Here θ refers to the weights and biases of the neural network NN. A lower physics informed loss indicates that the neural network is actually satisfying the PDE, and thus predicting the field more accurately. We re-emphasize that there is no data term in $f(\varepsilon; \theta)$, which simplifies the neural network training process. Furthermore, we believe that it mitigates the accumulation of error in the gradients during the inverse design process observed by Chen et. al.[47]. Figure 1 outlines the general strategy for building the proxy model. During each epoch, 10 (batch size) dielectric distributions consisting of rectangular pillars of height $h = 0.6$ μm with dielectric constant 4 (corresponding to SiN), are generated from 11 random pillar half-widths per batch. The operation wavelength is $\lambda = 0.633$ μm. The neural network architecture chosen is a UNET, shown in Fig. 1a and **b**, due to previously reported good performance with scattering problems[47]. The model is trained for 5×10^5 epochs using the ADAM optimizer[50] with a learning rate set to 5×10^{-4}. The final residual of the fields predicted by the neural network are of the order of ~0.5, compared to the numerical residual produced by FDFD which is on the order of 10^{-16}. Although there is a large difference, in the next section we show that this still produces a simulator which is capable of outperforming the LPA when optimizing the efficiency of a metalens. Figure 2**a** shows an example of a field predicted from a random set of pillars by the neural network, by a 2D FDFD code, and their difference, showing good qualitative match. A more quantitative measure of the errors is shown in Fig. 2**b**, where we show the point-wise error probability density functions for the relative error between the complex fields predicted by FDFD and that predicted by the neural network and the field predicted under LPA, and the absolute error between pillar-wise average transmission
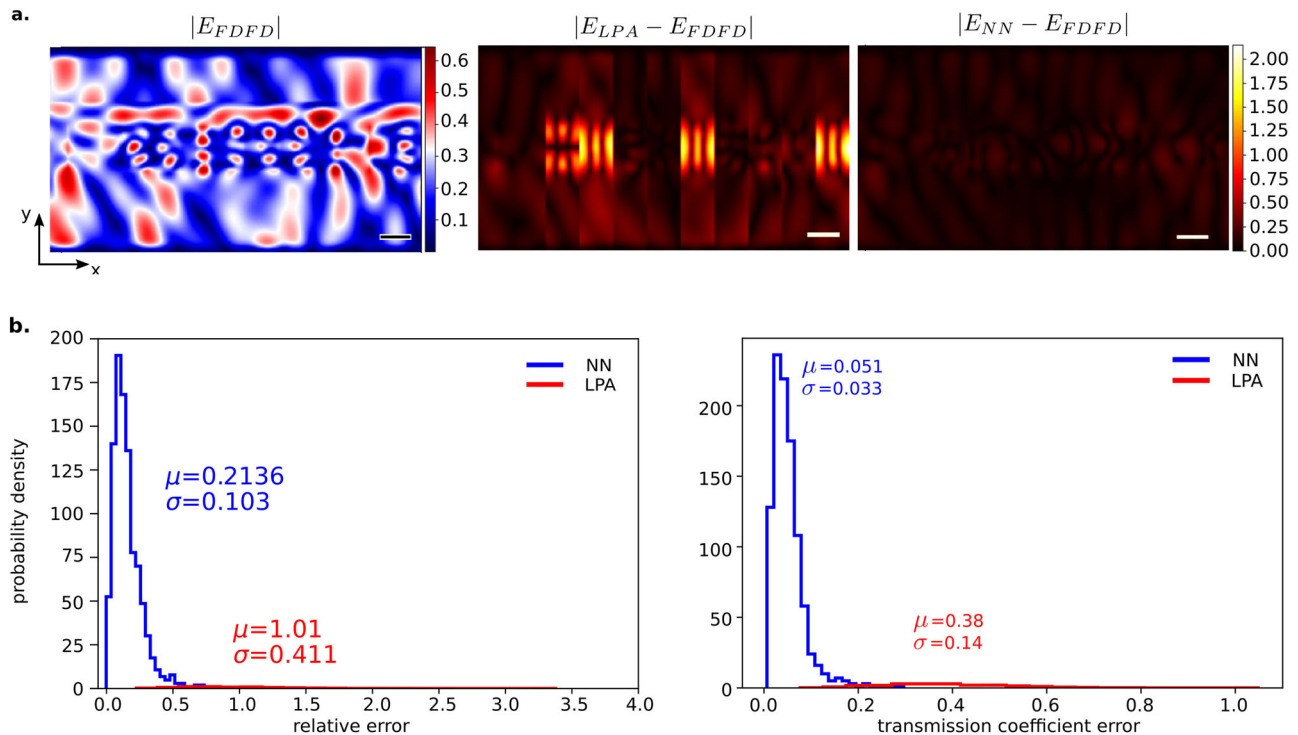
Fig. 2 Neural network performance analysis. a An example field. (left) is the absolute value of the ground truth finite difference frequency domain (FDFD) simulated field, (middle) difference between the true value of the field and the field produced by using the local phase approximation (LPA), and (right) is the absolute difference between the neural network field and the FDFD field. **b** Comparison between the performance of the proposed neural network and LPA methods. (left) Shows the relative error between FDFD predicted fields and the fields predicted by LPA. (right) Error comparisons between the transmission coefficients predicted by LPA and the neural net.

coefficients. See Supplementary Note 3 for a more detailed description of the pillar wise transmission coefficient error. The relative error is expressed as:

$$\frac{\left\|E_{\mathrm{approx}} - E_{\mathrm{FDFD}}\right\|_2^2}{\left\|E_{\mathrm{FDFD}}\right\|_2^2}. \tag{6}$$

For the PINN, E_{approx} is the field predicted from a set of 11 pillars. For the LPA, E_{approx} is fields predicted from the same set of pillars, and then stitched together over the same region. See Supplementary Fig. 2 for a visual explanation. The mean expected relative error for the neural network is $\mu = 0.21$ with a standard deviation of $\sigma = 0.103$. When using the LPA over the same region, we get a mean relative error of $\mu = 1.01$ with a standard deviation of 0.411. Thus, based on the relative field error, our method is $4.8\times$ more accurate than the LPA. For the pillar-wise transmission coefficient error, we get an expected error of $\mu = 0.051$ for the neural network with a standard deviation of $\sigma = 0.033$ and for the LPA method we get an expected error of $\mu = 0.38$ with a standard deviation of 0.14. Thus, based on the transmission coefficient error, our method is $7.2\times$ more accurate than the LPA.

Device optimization. The optimization process based on automatic differentiation functionality of PyTorch for large area meta-optics is outlined in Fig. 3. The forward problem is solved via a pre-trained PINN. Since the input into the neural net is a meshed grid of pillars, a differentiable map from pillar half-widths (3a.) to meshed geometries (3b) must be generated. This is achieved by generating Gaussian functions centered around pillar centers, with standard deviations of pillar half-widths in the x dimension, and pillar height in the y dimension, and then using a modified

softmax function to transform the Gaussians into rectangles with slightly rounded edges, making them differentiable via automatic differentiation (see Supplementary Note 4 and Supplementary Fig. 3). The meshed structures are fed into two separate neural networks that have been pre-trained to predict the complex electric field (3c.). The fields are then stitched together with regions of the outer half-widths overlapping. The total field is then propagated using the angular spectrum method (3d). The propagated field is used to calculate the FOM f (3e.)from Eq. (5). We use automatic differentiation to compute the gradients of the FOM with respect to the input half-widths $\nabla_{\overrightarrow{r}} f$, and iteratively update them with the ADAM optimizer[50].

Results

We used the PINN surrogate model to optimize 9 different lenses, all with 1 mm aperture, with focal lengths ranging from 250 to 1500 μm in increments of 250 μm. The minimum feature size is set to 75 *nm*, to ensure fabricability. To compare our optimization approach, we also generated lenses according to the hyperboloid phase equation:

$$\phi(x, y) = \frac{2\pi}{\lambda}\left(\sqrt{x^2 + F^2} - F\right) \tag{7}$$

The phase is implemented under LPA using SiN (refractive index 2), a wavelength of 0.633 μm, and periodicity of $p = 0.443$ μm (see Supplementary Fig. 9). We then optimize the lens employing our PINN to increase the intensity at the focal spot, i.e., the FOM is given by:

$$f = |E(x = 0, z = F)|^2 \tag{8}$$

Figure 4**a** and **b** show the intensity profile of a forward designed and optimized lens with $F = 500$ μm focal length. Figure 4**c** shows the normalized intensity slice at the focal spot of both lenses. As
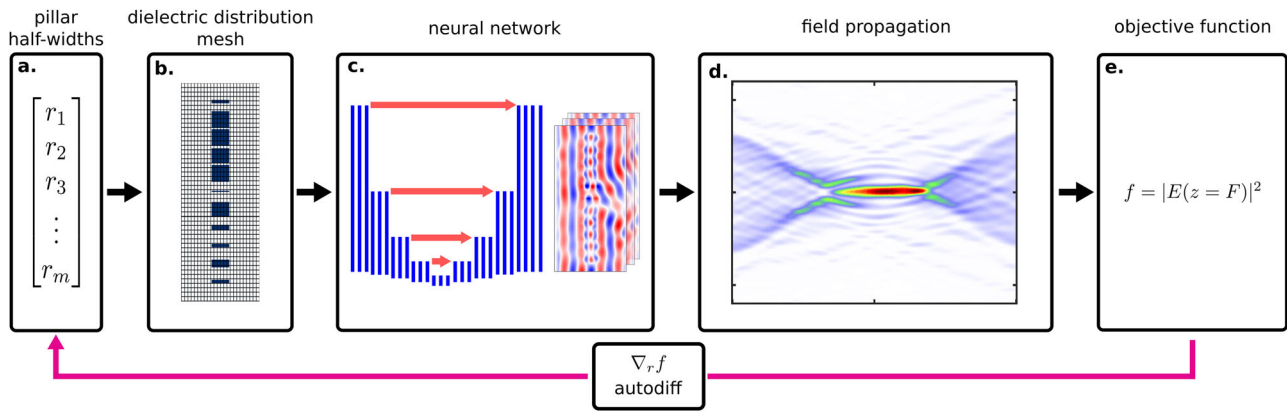
Fig. 3 Optimization strategy of 2D meta-optics with physics informed neural networks (PINNs). a We start with a vector, which contains a list of all pillar half-widths, characterizing the meta-optic. These half-widths are then batched into groups of 11 with an overlap of 1 pillar on each side (see Supplementary Note 2 and Supplementary Fig. 1 for notes on the computational domain set up). The choice of 11 pillars was made based on the GPU memory required to train the PINN. **b** The half-widths are meshed into dielectric distributions which get fed into the neural network. **c** The neural network predicts patches of fields which are then stitched together, and **d** propagated via the angular spectrum method. **e** The objective function is formed from the resulting field, and backpropagated using PyTorch's automatic differentiation functionality to update the initial radius distribution.
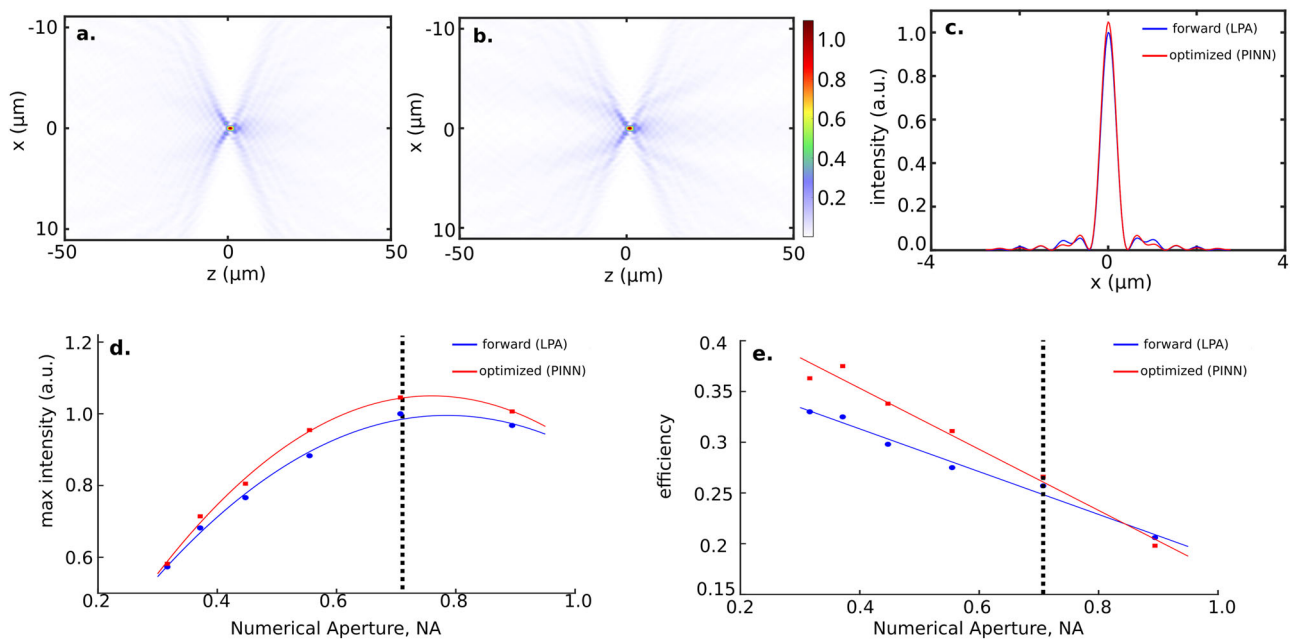

Fig. 4 Efficiency and intensity sweeps of forward designed lenses and optimized lenses. a Focal spot intensity profile of a forward designed lens with focal length $F = 500 \mu m$. **b** Focal spot intensity profile of an optimized lens. **c** Slices of intensity profiles for both lenses. The intensity was normalized such that the maximum intensity of the forward designed lens is 1. The theoretical performance improvement is ~3%. **d** Maximum intensity at the focal spot vs lens numerical aperture (NA). Intensities are normalized such that the maximum of the largest forward designed intensity is set to 1. **e** Theoretically computed efficiencies of the lenses vs NA. The solid lines are visual aids for the trend and do not correspond to a theoretical prediction.

seen in Fig. 4**d** the maximum intensities at the focal spots improve in every case. Figure 4**e** shows that the efficiency improves in all except for the lens with the highest NA. We also find a trend that the improvement in the maximum intensity of the inverse-designed meta-lens over the forward-design meta-lens increases with increasing NA. As with higher NA, the phase gradient becomes larger, we expect the LPA to be a worse approximation. Interpreting the efficiency improvement is more convoluted. We defined the efficiency as the ratio of the light energy inside a circle of radius of three times the full width half maximum (FWHM) at the focal spot over the total energy in the focal plane. With increasing NA, the FWHM decreases, making the efficiency improvement lower with increased NA. For the highest NA, the FWHM of the inverse-designed meta-lens is

significantly lower than the forward-designed meta-lens, making the efficiency lower.

We validated our designs by fabricating and experimentally testing the meta-lenses using a microscope (details of fabrication and characterization in Supplementary Note 1.1 and Supplementary Note 1.2). Figure 5 shows an example of the inverse optimized device. Figure 5**a–c** shows the scanning electron micrographs (SEMs) of the fabricated optimized lens with focal length $F = 500 \mu m$. Figure 5**d** shows the distribution of the dielectric pillar half-widths of the same forward and optimized lens. signifying the two designs are very different. Figure 5**e** shows the focal spot intensities of the lenses integrated over a $r = 3 \times$ FWHM region at the focal spot, which yields a quantitative value to compare the lens efficiency[51] among different devices.
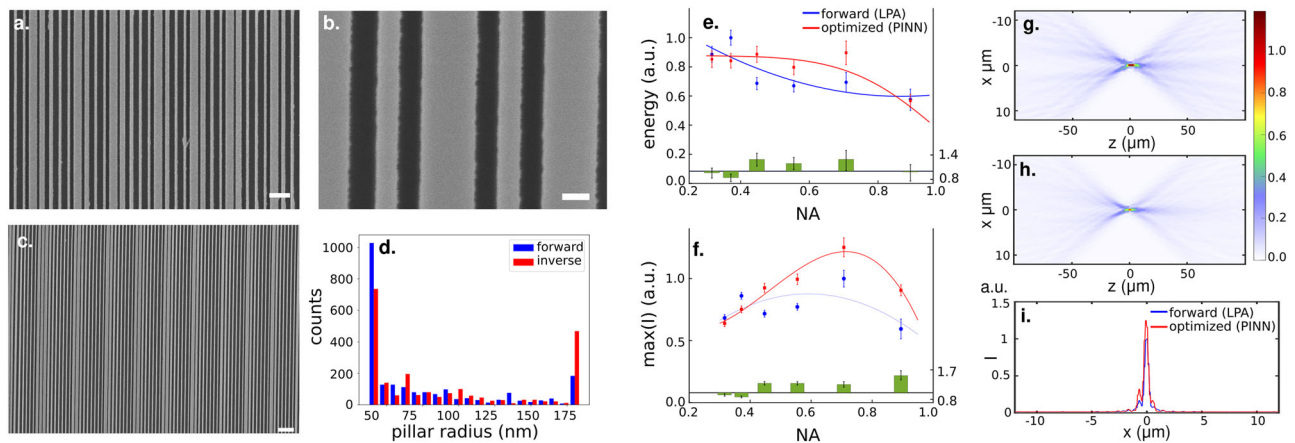
Fig. 5 Experimental results. a–c Scanning Electron Microscope (SEM) images of the fabricated SiN meta-lens with focal length $F = 500\mu m$. The scale bars correspond to 1 μm, 0.1 μm, and 1 μm, respectively. **d** Counts of pillar half-widths of the forward and inverse designed lens. **e** Measured intensity contained in the region given by 3 × FWHM of the focal spot vs lens numerical aperture. The units are normalized to the largest intensity integral of the forward design. Error bars are standard deviations of energy computed at 1216 data points (dimension of the sensor). **f** Maximum intensity at the focal spot. The inverse designed lenses outperform the forward designed lenses for NA > 0.44. The lines are visual aids and not fits to a theoretical model. Units are normalized to the largest intensity of the forward designed lens. In the NA = 0.9 ($F = 250\mu m$) case an improvement of 53% is observed. The error bars are standard deviations of intensity at the focal plane. **g** Experimentally measured field intensities of the forward designed lens and **h** of the inverse designed lens. **i** Intensity slice at the focal spot. The intensities are normalized such that the maximum intensity of the forward designed lens is 1. The intensity of the inverse designed lens focal spot is 1.25.

Figure 5**f** plots the maximum intensity plot as a function of the lens NA. For optimized lenses with NA > 0.44, we see improvements of more than 25%, with a maximum improvement of 53% for the NA = 0.9 lens. The experimentally determined intensity integral, which is analogous to the efficiency of a lens, on has improvements of more than 18% in all cases except for the NA=0.9 case. This is because the FWHM of the optimized lens at the NA=0.9 case is actually smaller than the FWHM of the forward designed lens, leading to a smaller integration area when computing the energy. We note that, while a quantitative match between the experiment and design is not obtained, we did observe a similar trend in terms of improved intensity and efficiency as predicted by the theory. Fig. 5**g** shows experimentally measured field profiles of the forward designed $F = 500\mu m$ meta-lens. Figure 5**h** shows the same for an optimized lens. Figure 5**i** is the slice of the focal spot intensity profile along the $z = F$ plane. In all these figures, the intensity is normalized such that the maximum intensity of the forward designed lens is 1.

Discussion
We have developed a PINN to use as a proxy surrogate model for simulating the full Maxwell's equations to design dielectric meta-optics. We used the PINN to optimize pillar half-widths to maximize the intensity at the focal spot of 1 mm aperture cylindrical meta-lenses at 633nm. We demonstrated experimental improvements of the maximum intensity of the lenses up to 53%. We also want to note that this method was useful for the inverse design of extended depth of focus lenses[10] (see Supplementary Note 7 and Supplementary Fig. 6). This model did not use the LPA, but simulated meta-atoms by splitting up the device into chunks with overlapping boundaries, and stitching the chunks together to approximate the full field response. We emphasize that FDFD simulations were never carried out to train the PINN, and we only minimized the residual of the PDE itself to train the network. The PINN training took approximately 2 hours on our machine. In our studies, this method provided approximately a 3-5x speedup (see Supplementary Figure 8 and Supplementary Table 1) over conventional FDFD with overlapping boundary conditions, and was

much simpler to use as a forward simulator for optimization problems since it can be used as a simple map from ϵ to E-field with gradients computed by automatic differentiation.

We would also like to note that the theoretical intensity and efficiency improvements are smaller than their experimental counterparts. While we do not have a clear explanation for this discrepancy, the theoretical and experimental trends in lens improvement are similar. One hypothesis could be that the the inverse designed lenses may be more tolerant to fabrication imperfections. However, randomly changing the scatterers in our meta-optics by 10% did not give a similar enhancement. As such this aspect of the quantitative match between experiment and theory remains an open problem and further studies will be needed. We would like to point out however the importance of experimentally verifying inverse design methodologies, since in our studies we used open source codes that produce reasonable looking results, but are not experimentally accurate (see Supplementary Note 8 and Supplementary Fig. 7). The inverse design solution we introduced in this paper can be integrated into various computationally intensive tasks which require mate-optical inverse design such as the end-to-end optimization of computational imaging systems and the design of optical neural networks[52,53]. It is worth noting, however, that this method is not a general numerical solver. It is limited to predicting electromagnetic field responses from fixed source, material, and boundary parameters. Source type and k-vector, dielectric constant, geometry type (rectangular pillar of fixed height), and boundary conditions must all remain constant for this method to work. If any of these parameters are modified, the PINN must be retrained. Furthermore, the method we presented was only implemented under a 2D approximation. Extending this method to 3 dimensions would take significant effort due to the fact that the electric field $\mathcal{E}$ could no longer be treated as a scalar field, and the full vector nature would have to be modeled. On a $n \times n$ grid in 2D, the Maxwell operator $[\nabla^2 + \omega^2 \epsilon]$ results in a $n^2 \times n^2$ matrix, while for a $n \times n \times n$ 3D grid the Maxwell operator $[\nabla \times \nabla \times - \omega^2 \epsilon]$ result in $9n^3 \times 9n^3$ square matrices due to the additional 2 vector field components that must be modeled. However, these operators are sparse with a small number of

nonzero elements that scale as $\sim 38n^3$ in 3D, making small problems still manageable. The other problem with generalizing this method to 3D is the large null-space of the $\nabla \times \nabla \times$ operator which results in slow convergence of numerical methods[54,55]. It is highly likely that this could also affect the training of the PINN, and require regularization or preconditioning which deflates the null space of this operator to properly converge onto a solution. On the other hand, in this work we showed that machine-precision numerical accuracy of numerical solvers may be not be needed for inverse design methods with FDFD. Solvers could be sped up by relaxing the relative error tolerance, such that iterative solvers converge quicker for predicting the forward and adjoint problems. Another interesting aspect will be to understand the optimal PINN to model the meta-optics, and if we can identify a relationship between the number of trainable parameters and size of the problem we are solving. In future work we aim to explore these options.

Data availability

Code availability

References

1. Şumnu, A., Güzelbey, İ. H. & Öğücü, O. Aerodynamic shape optimization of a missile using a multiobjective genetic algorithm. *Int. J. Aerospace Eng.* **2020**, 1528435 (2020).
2. Molesky, S. et al. Inverse design in nanophotonics. *Nat. Photo.* **12**, 659–670 (2018).
3. Yee, K. Numerical solution of initial boundary value problems involving maxwell's equations in isotropic media. *IEEE Trans. Antennas Propagation* **14**, 302–307 (1966).
4. Rumpf, R. C., Garcia, C. R., Berry, E. A. & Barton, J. H. Finite-difference frequency-domain algorithm for modeling electromagnetic scattering from general anisotropic objects. *Prog. Electromag. Res. B* **61**, 55–67 (2014).
5. Shin, W. *3D Finite-Difference Frequency-Domain Method for Plasmonics and Nanophotonics*. Ph.D. thesis https://www.proquest.com/dissertations-theses/3d-finite-difference-frequency-domain-method/docview/2463202947/se-2 (2013).
6. Zhan, A., Fryett, T. K., Colburn, S. & Majumdar, A. Inverse design of optical elements based on arrays of dielectric spheres. *Appl. Opt.* **57**, 1437–1446 (2018).
7. Piggott, A. Y. et al. Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer. *Nat. Photo.* **9**, 374–377 (2015).
8. Piggott, A. Y., Petykiewicz, J., Su, L. & Vučković, J. Fabrication-constrained nanophotonic inverse design. *Sci. Rep.* **7**, 1786 (2017).
9. Pestourie, R. et al. Inverse design of large-area metasurfaces. *Opt. Exp.* **26**, 33732–33747 (2018).
10. Bayati, E. et al. Inverse designed metalenses with extended depth of focus. *ACS Photo.* **7**, 873–878 (2020).
11. Zhelyeznyakov, M. V., Brunton, S. & Majumdar, A. Deep learning to accelerate scatterer-to-field mapping for inverse design of dielectric metasurfaces. *ACS Photo.* **8**, 481–488 (2021).
12. Elsawy, M. M. R. et al. Global optimization of metasurface designs using statistical learning methods. *Sci. Rep.* **9**, 17918 (2019).
13. Park, J. et al. Free-form optimization of nanophotonic devices: from classical methods to deep learning. *Nanophotonics* **11**, 1809–1845 (2022).
14. Zhelyeznyakov, M. V., Zhan, A. & Majumdar, A. Design and optimization of ellipsoid scatterer-based metasurfaces via the inverse t-matrix method. *OSA Cont.* **3**, 89–103 (2020).
15. Munley, C. et al. Inverse-designed meta-optics with spectral-spatial engineered response to mimic color perception. *Adv. Opt. Mater.* **10**, 2200734 (2022).
16. Chung, H. & Miller, O. D. High-na achromatic metalenses by inverse design. *Opt. Exp.* **28**, 6945–6965 (2020).
17. Bayati, E., Zhan, A., Colburn, S., Zhelyeznyakov, M. V. & Majumdar, A. Role of refractive index in metalens performance. *Appl. Opt.* **58**, 1460–1466 (2019).
18. Mansouree, M. et al. Multifunctional 2.5d metastructures enabled by adjoint optimization. *Optica* **7**, 77–84 (2020).
19. Zhan, A. et al. Controlling three-dimensional optical fields via inverse mie scattering. *Sci. Adv.* **5**, eaax4769 (2019).
20. Li, W. F. et al. Transcending shift-invariance in the paraxial regime via end-to-end inverse design of freeform nanophotonics. https://arxiv.org/abs/2302.01712 (2023).
21. Byrnes, S. J., Lenef, A., Aieta, F. & Capasso, F. Designing large, high-efficiency, high-numerical-aperture, transmissive meta-lenses for visible light. *Opt. Exp.* **24**, 5110–5124 (2016).
22. Spägele, C. et al. Multifunctional wide-angle optics and lasing based on supercell metasurfaces. *Nat. Commun.* **12**, 3787 (2021).
23. Hsu, L., Dupré, M., Ndao, A., Yellowhair, J. & Kanté, B. Local phase method for designing and optimizing metasurface devices. *Opt. Exp.* **25**, 24974–24982 (2017).
24. Phan, T. et al. High-efficiency, large-area, topology-optimized metasurfaces. *Light Sci. Appl.* **8**, 48 (2019).
25. Lin, Z., Liu, V., Pestourie, R. & Johnson, S. G. Topology optimization of freeform large-area metasurfaces. *Opt. Exp.* **27**, 15765–15775 (2019).
26. Torfeh, M. & Arbabi, A. Modeling metasurfaces using discrete-space impulse response technique. https://arxiv.org/abs/2003.06683 (2020).
27. Skarda, J. et al. Low-overhead distribution strategy for simulation and optimization of large-area metasurfaces. *Comput. Mate.* **8**, 78 (2022).
28. Shlezinger, N., Whang, J., Eldar, Y. C. & Dimakis, A. G. Model-based deep learning. https://arxiv.org/abs/2012.08405 (2020).
29. Malkiel, I. et al. Plasmonic nanostructure design and characterization via deep learning. *Light: Sci. Appl.* **7**, 60 (2018).
30. Li, X., Shu, J., Gu, W. & Gao, L. Deep neural network for plasmonic sensor modeling. *Opt. Mater. Exp.* **9**, 3857–3862 (2019).
31. Peurifoy, J. et al. Nanophotonic particle simulation and inverse design using artificial neural networks. *Sci. Adv.* **4** (2018). https://advances.sciencemag.org/content/4/6/eaar4206. https://advances.sciencemag.org/content/4/6/eaar4206.full.pdf.
32. Kiarashinejad, Y. et al. Knowledge discovery in nanophotonics using geometric deep learning. *Adv. Intell. Syst.* **2**, 1900132 (2019).
33. So, S., Badloe, T., Noh, J., Bravo-Abad, J. & Rho, J. Deep learning enabled inverse design in nanophotonics. *Nanophotonics* **9**, 1041–1057 (2020).
34. Liu, D., Tan, Y., Khoram, E. & Yu, Z. Training deep neural networks for the inverse design of nanophotonic structures. *ACS Photo.* **5**, 1365–1369 (2018).
35. Gao, L., Li, X., Liu, D., Wang, L. & Yu, Z. A bidirectional deep neural network for accurate silicon color design. *Adv. Mater.* **31**, 1905467 (2019).
36. An, S. et al. A deep learning approach for objective-driven all-dielectric metasurface design. *ACS Photo.* **6**, 3196–3207 (2019).
37. Raissi, M., Perdikaris, P. & Karniadakis, G. 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).
38. Lu, L., Meng, X., Mao, Z. & Karniadakis, G. E. Deepxde: A deep learning library for solving differential equations. *SIAM Rev.* **63**, 208–228 (2021).
39. Karumuri, S., Tripathy, R., Bilionis, I. & Panchal, J. Simulator-free solution of high-dimensional stochastic elliptic partial differential equations using deep neural networks. *J. Comput. Phys.* **404**, 109120 (2020).
40. Raissi, M., Yazdani, A. & Karniadakis, G. E. Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations. *Science* **367**, 1026–1030 (2020).
41. Zhu, Y., Zabaras, N., Koutsourelakis, P.-S. & Perdikaris, P. Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data. *J. Comput. Phys.* **394**, 56–81 (2019).
42. Tartakovsky, A., Marrero, C., Perdikaris, P., Tartakovsky, G. & Barajas-Solano, D. Physics-informed deep neural networks for learning parameters and constitutive relationships in subsurface flow problems. *Water Res. Res.* **56**, e2019WR026731 (2020).
43. Yazdani, A., Lu, L., Raissi, M. & Karniadakis, G. E. Systems biology informed deep learning for inferring parameters and hidden dynamics. *PLoS Comput. Biol* **16**, e1007575 (2020).
44. Zhang, D., Lu, L., Guo, L. & Karniadakis, G. E. Quantifying total uncertainty in physics-informed neural networks for solving forward and inverse stochastic problems. *J. Comput. Phys.* **397**, 108850 (2019).
45. Chen, Y., Lu, L., Karniadakis, G. E. & Dal Negro, L. Physics-informed neural networks for inverse problems in nano-optics and metamaterials. *Opt Exp.* **28**, 11618–11633 (2020).
46. Lu, L. et al. Physics-informed neural networks with hard constraints for inverse design. *SIAM J. Sci. Comput.* **43**, B1105–B1132 (2021).

47. Chen, M. et al. High speed simulation and freeform optimization of nanophotonic devices with physics-augmented deep learning. *ACS Photonics.* https://doi.org/10.1021/acsphotonics.2c00876 (2022).

48. Valantinas, L. & Vettenburg, T. A physics-defined recurrent neural network to compute coherent light wave scattering on the millimetre scale. https://arxiv.org/abs/2208.01118 (2022).

49. Hughes, T. W., Minkov, M., Williamson, I. A. D. & Fan, S. Adjoint method and inverse design for nonlinear nanophotonic devices. *ACS Photo.* 5, 4781–4787 (2018).

50. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).

51. Arbabi, A., Horie, Y., Ball, A. J., Bagheri, M. & Faraon, A. Subwavelength-thick lenses with high numerical apertures and large efficiency based on high-contrast transmitarrays. *Nat. Commun.* 6, 7069 (2015).

52. Tseng, E. et al. Neural nano-optics for high-quality thin lens imaging. *Nat. Commun.* 12, 6493 (2021).

53. Lin, X. et al. All-optical machine learning using diffractive deep neural networks. *Science* 361, 1004–1008 (2018).

54. Newman, G. A. & Alumbaugh, D. L. Three-dimensional induction logging problems, Part 2: A finite-difference solution. *Geophysics* 67, 484–491 (2002).

55. Newman, G. & Weiss, C. Electromagnetic induction in a generalized 3d anisotropic earth, part 2: The lin preconditioner. *Geophysics* 68, 922–30 (2003).

Acknowledgements

Author contributions

M.V.Z. conceptualized the project. M.V.Z. designed the methodology, wrote the software, trained the machine learning model, and performed the data analysis. J.F. manufactured the metasurfaces and experimentally validated them. A.W.S. validated metasurface designs in simulation. J.N. assisted with validating metasurface designs in simulation. J.R. provided compute resources for training the machine learning model. S.B. and A.M. supervised the project. Initial version of manuscript was written by M.V.Z.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44172-023-00107-x.

Correspondence and requests for materials should be addressed to Maksym Zhelyeznyakov or Arka Majumdar.

Peer review information : *Communications Engineering* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Rosamund Daw, Mengying Su. A peer review file is available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.