PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Machine learning for the design of nanomaterials

Blanchard-Dionne, André-Pierre, Martin, Olivier J.

André-Pierre Blanchard-Dionne, Olivier J. F. Martin, "Machine learning for the design of nanomaterials," Proc. SPIE 11462, Plasmonics: Design, Materials, Fabrication, Characterization, and Applications XVIII, 114621C (20 August 2020); doi: 10.1117/12.2568471

SPIE.

Event: SPIE Nanoscience + Engineering, 2020, Online Only

Invited Paper

Machine Learning for the Design of Nanomaterials

André-Pierre Blanchard-Dionne[a] and Olivier J.F. Martin[b]

[a,b]Ecole Polytechnique Federale de Lausanne, Switzerland

ABSTRACT

In this article we review some machine learning methods for the design of nanomaterials. The first part will discuss how to use neural network to build a predictive model of the optical properties of a certain material or structure. The second part is dedicated to the optimization and reverse engineering of an optical material using generative networks.

Keywords: Machine Learning, Nanomaterials, Optical Materials, Plasmonics, Metamaterials

1. INTRODUCTION

Machine learning has recently emerged as a new tool for nanophotonics to improve data analysis[1,2] both at the simulation and experimental levels. At the experimental level, neural networks can enhance our interpretation of data and understanding of patterns; it has found applications for refractive index sensors,[3] SERS experiments[4] or for data storage devices.[5] On the simulation level, machine learning can draw relationships between the parameters of an object and its optical properties and can thus serve as a predicting model[6] of the material studied. This model can in turn be used in a more complex reverse engineering algorithm[7,8] to find a solution to an optimization problem. The more straightforward approach is to use a reverse network linking the optical property to the parameters of the object, as it has been shown for reflection coating,[9] waveguides,[10] power splitters,[11] nanostructures,[12] nanoparticles[13] or chiral metamaterials.[14] Another method widely implemented is generative algorithms, which create solutions targeted towards a certain property, a method very well suited for geometry-based problems. It has found applications for metasurfaces,[15–19] nanostructures,[20] metagratings,[21] thermal emitters[22] chiral metasurfaces,[23] photonic crystals,[24,25] power splitters[26] and optical cloaks.[27] Another approach uses reinforcement learning to guide the process of selecting the optimal parameters of an object, and has been demonstrated for color generation.[28]

In what follows we introduce some basic elements of neural networks as well as strategies to improve their efficiency. We will consider forward predictive networks in a first part and how we can include physics-based equations to improve their generalization capabilities.[29] In a second part, we will present how we can use a feedback loop for a generative adversarial network in order to optimize the geometry of an all-dielectric optical cloak.[27]

2. FORWARD PREDICTIVE NETWORK

2.1 Fully connected Networks

Fully connected networks (FCN) are meant to draw a relationship between an object's parameters and its optical properties. Once trained, the network constitutes a synthetic function that takes a series of parameters as inputs and calculates an output via the weights of every node of the network. The most simple case of a FCN is depicted in Figure 1 for the prediction of the transmission spectrum of a metallic grating. It has the period and width of a nanoslit array as inputs, the spectrum as the output and has 3 hidden layers. In this case of supervised training, a dataset of previously collected spectra will be used to update the weights of the hidden layers in order to build an optimal model. The spectra were obtained using coupled mode theory.[30–32]

This type of network, when trained with a sufficient number of example and with adequate activation function and number of nodes, performs relatively well for parameters close to those used for training, as will be shown in the next section. However the performance decreases drastically with very different input parameters. For this reason, different strategies are employed to improve the capabilities of the network to generalize its predictions.

Further author information: (Send correspondence to Andre-Pierre Blanchard-Dionne: E-mail: andre-pierre.blanchard-dionne@epfl.ch

Plasmonics: Design, Materials, Fabrication, Characterization, and Applications XVIII,
edited by Din Ping Tsai, Takuo Tanaka, Proc. of SPIE Vol. 11462, 114621C
 · doi: 10.1117/12.2568471

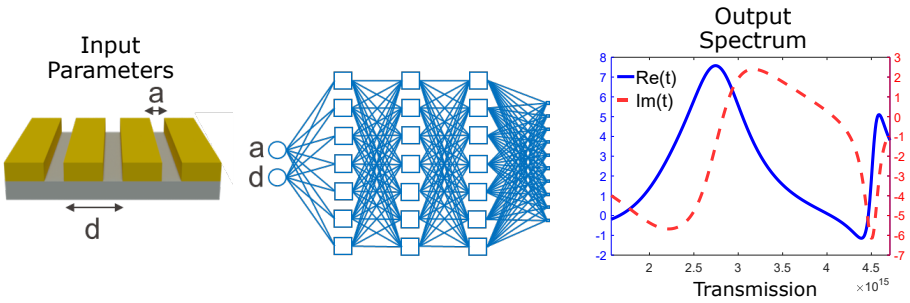

Figure 1. A fully connected network for the prediction of the optical spectrum of a metallic nano-grating.

2.2 Networks with physics-based equations

In this section we show how we can improve the performance of neural networks when we include in their structure a set a fundamental physics-based equations.[29] Asa physical system, we consider a nanoslit array and use the fact that its transmission is determined by the resonance states of the structure, we can use the basic Lorentzian oscillator equations and a background factor to fully describe the spectrum. The network used is represented in Figure 2. The parameters are connected to a series of hidden layers which give as output the parameters of a

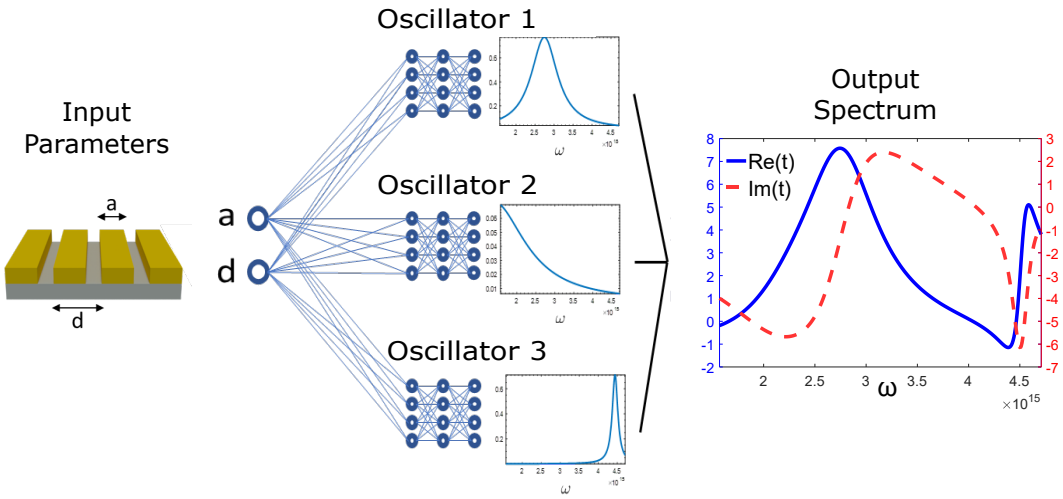

Figure 2. Neural Network including Lorentz oscillators. The output spectrum is the sum of the response of each oscillator.

Lorentz oscillator ω, γ, c and θ respectively the resonant central frequency, the losses, the amplitude and phase of the oscillator. The actual output of each oscillator is thus given through the general equation :

$$l(\omega) = \frac{ce^{i\phi}}{\hbar(\omega - \omega_0) + i\gamma}, \tag{1}$$

and the total transmission will be given by the summation of the response of each oscillator.

In Figure 3 the transmission obtained using this method is presented alongside the response of of each individual oscillator. We can observe that the two main resonance peaks of (a) are each well represented by a single oscillator, the blue and yellow curves in (b). The advantage of the method can be clearly observed in the following Figure 4, where the relationship between the inverse of the central frequency $2\pi c/\omega_0$, or the resonant wavelength, and the losses γ are plotted as a function of the period d. According to the physics of the resonance[30] the former should have a linear behaviour with the period while the latter should have an inverse behaviour. This is observed even for values of d which falls outside the range of the training set, meaning the the network was able to grasp the physics of the studied material and extrapolate it.

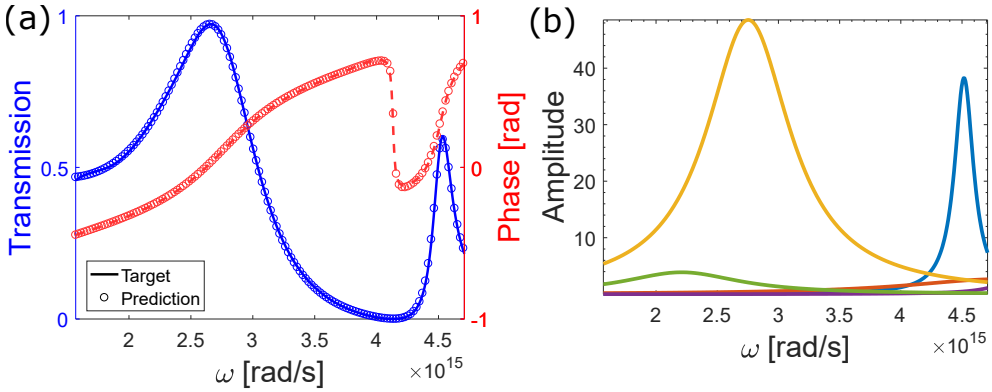

Figure 3. (a) Predicted and target spectrum for the model using oscillator equations for a nanoslit array of $d = 266$ nm and a slit width of $a = 66$ nm. (b) Output of each individual oscillator.

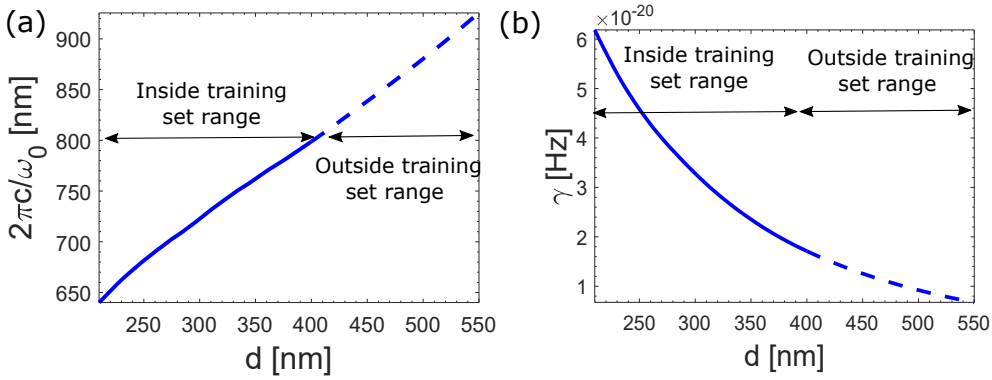

Figure 4. Relationship for (a) the resonant wavelength $\lambda_0 2\pi c/\omega_0$ and (b) the losses γ as a function of the periodicity of the system.

The fact that the spectrum follows a phenomenological behaviour leads to a much more robust and accurate model, as can be seen in Figure 5 where the predicted spectrum for a FCN and the model with oscillators are presented for parameters outside the training set range.

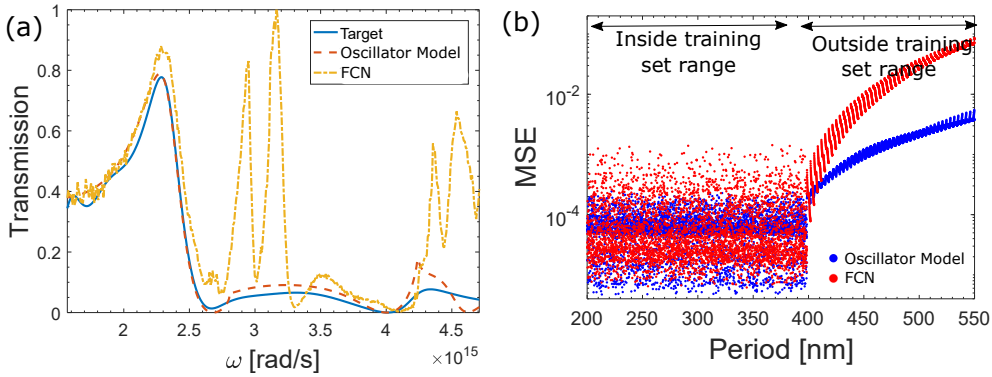

Figure 5. (a) Predicted spectrum with the oscillator model (red dashed line) and fully connected network (yellow dashed-dotted line) compared to the target (blue line). (b) Mean squared error of both models for values of the periodicity inside and outside the training range

The FCN (yellow dashed curve) shows some spurious peaks, which could be expected since the fitting of the spectrum in such a point-by-point method is more subject to inconsistencies or overfitting of the model. The model with oscillators (red dashed curve) follows the behaviour of the individual oscillator and thus gives a clearly better match with the target curve. We can have a global idea of the improvement of this model by

looking at Figure 5(b), which plots the mean squared error of the predicted spectrum for different values of the period. Both networks show good performances of 1.02×10^{-4} and 6.01×10^{-5} inside the range of the training set, but reach in average 0.031 and 0.0016 outside of it, for an improvement of a factor of 20 for the model with oscillators.

3. GENERATIVE NETWORK FOR REVERSE ENGINEERING

Generative networks were first created mostly for computer vision applications. They have the ability to manipulate images and create realistic new images of a certain subject.[33] They can be a great tool for geometry-based optimization of nanomaterials since they can efficiently generate new design. Furthermore they can also be trained for generating objects that satisfy a certain optimization goal. In this section we will present a generative adversarial network for the optimization of an optical cloak.[27]

3.1 Convolution Network

We use a convolution network to create a predictive forward network with an image as input. The convolution operation aims at recognizing features in the image. For this study we optimize the geometry of a cloaking shell as illustrated in Figure 6(a), which tries to conceal a perfectly reflective object at R1. The shell has a uniform isotropic dielectric constant of $\epsilon = 2$ and the background is air with $\epsilon = 1$. The convolution network (Figure 6(b)) takes a binary image of the shell as input and predicts its scattering coefficient Ψ_p as output. 10 000 random shells were randomly generated and their scattering coefficients were computed with the finite-element method to create the training dataset.

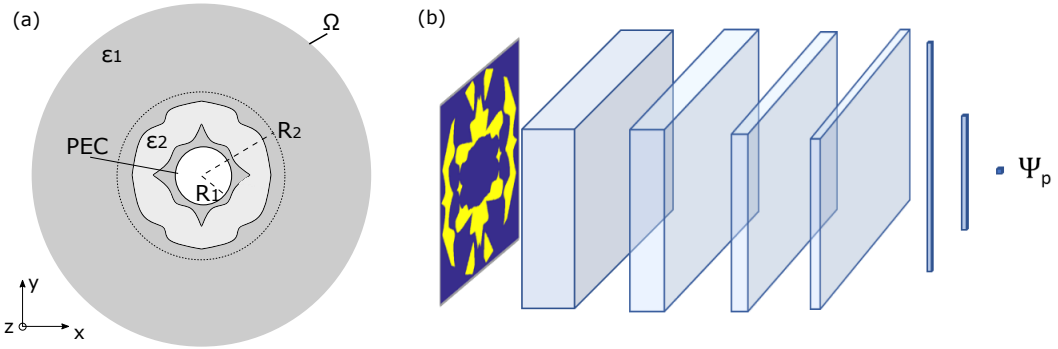

Figure 6. (a) Geometry of the problem for cloaking with a uniform isotropic dielectric shell. (b)

Once the convolution network has been trained, it can be included in a generative adversarial network (GAN) in order to generate new prototypes of the shell. This network is made of the generator, a convolution upsampling network which takes noise input and generate images, and the discriminator, a convolution network which aims a distinguishing which images are "real" (from the dataset) and which are "false" (from the generator) by having this probability p as output. To this branched network we add the convolution network for which the losses are added to the generator. The GAN in this setup (see Figure 7) accomplishes 2 tasks: it creates configurations of the shells which resemble those of the dataset (avoiding noisy and random configurations) and which minimize the scattering coefficient Ψ.

The training of the GAN is presented in Figure 8. The generator and discriminator usually reach an equilibrium state since their loss calculation is in opposition. The forward network in this case adds an additional loss to the generator, which shifts the losses of the generator towards higher values. Proper scaling of the loss factors must be done in order to generate adequate potential solutions. In Figure 8(b) we can observe the evolution of the generation of shell configurations for different epochs of the training. The first few examples are slightly more random but then tend towards prototypes which represent good solutions for cloaking.

In order to improve the solution search using the GAN, we implement a feedback algorithm which takes the best proposed solution by the GAN, calculates their actual scattering coefficent (and not the predicted one from

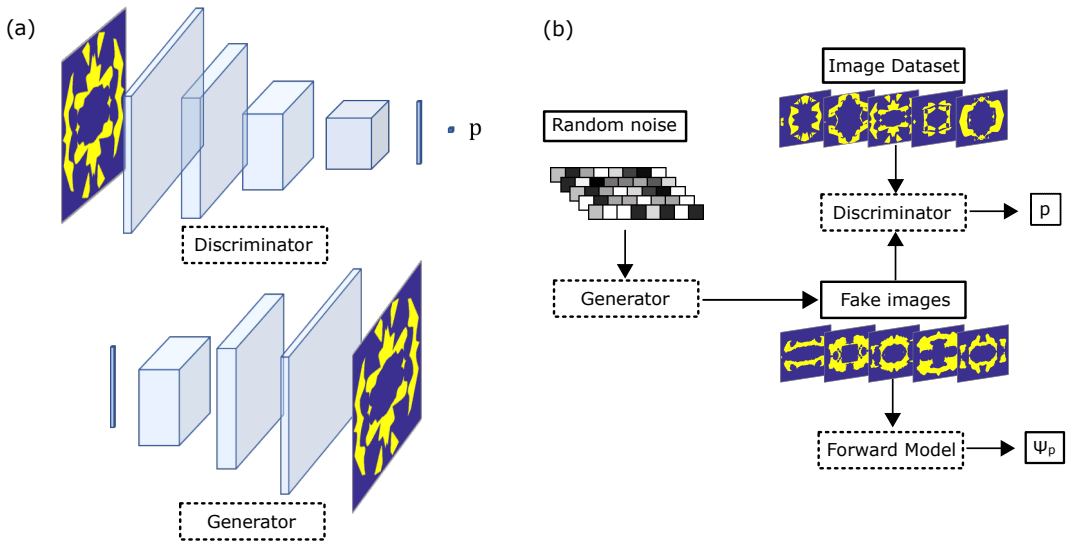

Figure 7. (a) The discriminator, a convolution network, and Generator, a transposed convolution network. (b) Structure of the generative adversarial network

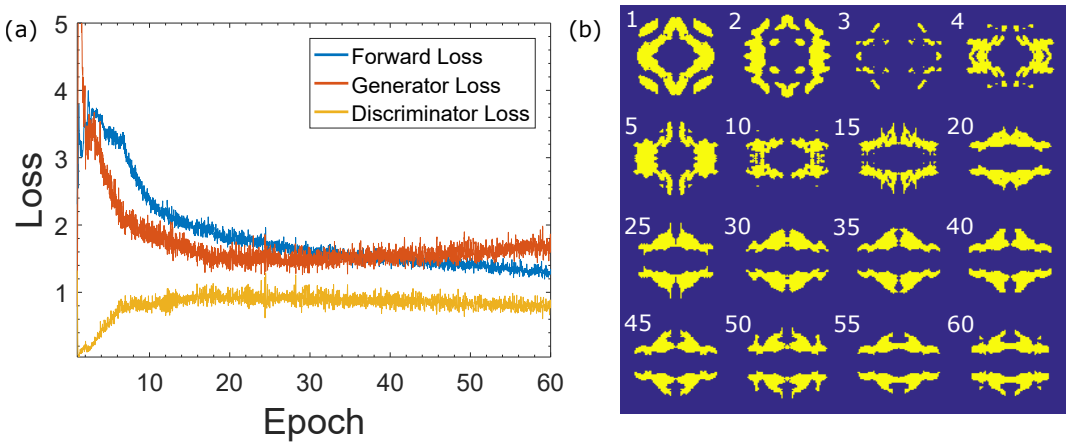

Figure 8. (a) Training of the GAN. (b) Evolution of the images created by the GAN during training.

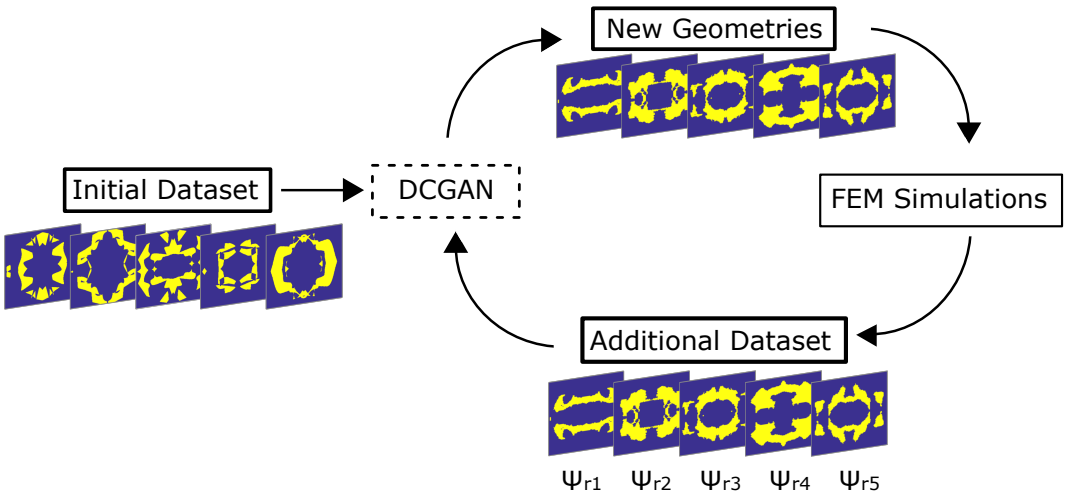

Figure 9. Feedback loop used to improve the solution search

the forward network), adds them to the dataset and then retrains the GAN for another iteration. This process is illustrated in Figure 9.

The effect of the retraining improves the solution search in 2 ways: First, it improves the accuracy of the forward network by including configurations of the cloaks that were deemed optimal and by retraining it with the actual scattering coefficient. This way the forward network can learn which optimal solutions were actually good, and which weren't, reorienting the solution search. Second, the new solutions populate the dataset with favorable configurations, which will influence the GAN into generating images that are similar. We can follow the improvement of the proposed solutions by the GAN in Figure 10, where the average value of the calculated scattering coefficient of the solutions found as well as the minimum value found (best candidate), are plotted for each iteration of the feedback loop of the GAN.

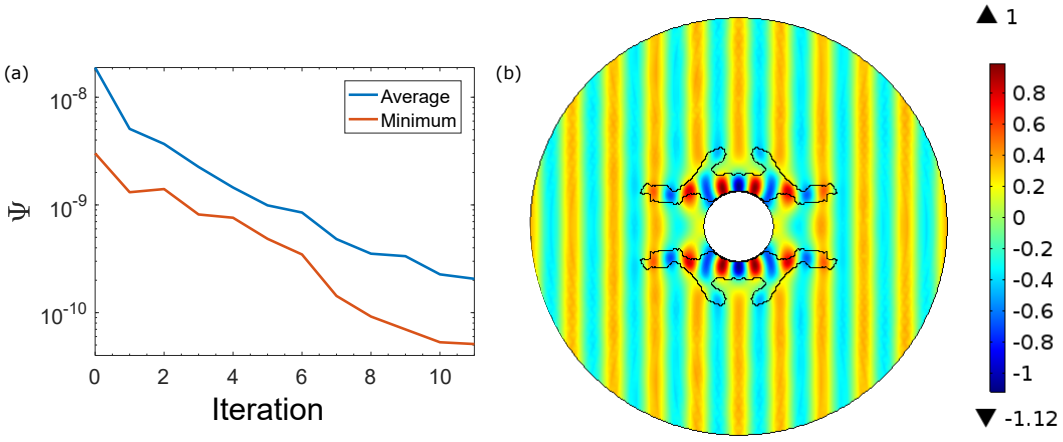

Figure 10. (a) Value of the average and minimal scattering coefficient for the 1000 answers generated at each training of the GAN. (b) Optimal configuration of the cloaking shell.

Each iteration shows an improvement until it reaches a minimum value of 5.11×10^{-11} W/m. This corresponds to a ratio of cloaking of 0.0089 when comparing to the value 5.77×10^{-9} W/m for the case without a shell, a result comparable to ones obtained using topology methods. The optimal configuration is presented in Figure 10 where we can see the bending of the electromagnetic wave around the object to conceal, a feature of cloaking shells.

4. CONCLUSION

In conclusion, we have shown how we can improve predictive forward network by giving them more structure to work with, in this case by including physics-based equations. We have also demonstrated how we can accomplish reverse engineering with a generative adversarial network, and how a feedback loop of this network can gradually improve the optimization process.

REFERENCES

[1] Yao, K., Unni, R., and Zheng, Y., "Intelligent nanophotonics: merging photonics and artificial intelligence at the nanoscale," *Nanophotonics* **8**(3), 339–366 (2019).

[2] Zhou, J., Huang, B., Yan, Z., and Bünzli, J.-C. G., "Emerging role of machine learning in light-matter interaction," *Light: Science & Applications* **8**(1), 1–7 (2019).

[3] Ballard, Z. S., Shir, D., Bhardwaj, A., Bazargan, S., Sathianathan, S., and Ozcan, A., "Computational sensing using low-cost and mobile plasmonic readers designed by machine learning," *ACS nano* **11**(2), 2266–2274 (2017).

[4] Banaei, N., Moshfegh, J., Mohseni-Kabir, A., Houghton, J. M., Sun, Y., and Kim, B., "Machine learning algorithms enhance the specificity of cancer biomarker detection using sers-based immunoassays in microfluidic chips," *RSC advances* **9**(4), 1859–1868 (2019).

[5] Wiecha, P. R., Lecestre, A., Mallet, N., and Larrieu, G., "Pushing the limits of optical information storage using deep learning," *Nature nanotechnology* **14**(3), 237–244 (2019).
[6] Sajedian, I., Kim, J., and Rho, J., "Predicting resonant properties of plasmonic structures by deep learning," *arXiv preprint arXiv:1805.00312* (2018).
[7] Sanchez-Lengeling, B. and Aspuru-Guzik, A., "Inverse molecular design using machine learning: Generative models for matter engineering," *Science* **361**(6400), 360–365 (2018).
[8] So, S., Badloe, T., Noh, J., Bravo-Abad, J., and Rho, J., "Deep learning enabled inverse design in nanophotonics," *Nanophotonics* **1**(ahead-of-print) (2020).
[9] Liu, D., Tan, Y., Khoram, E., and Yu, Z., "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics* **5**(4), 1365–1369 (2018).
[10] Zhang, T., Wang, J., Liu, Q., Zhou, J., Dai, J., Han, X., Zhou, Y., and Xu, K., "Efficient spectrum prediction and inverse design for plasmonic waveguide systems based on artificial neural networks," *Photonics Research* **7**(3), 368–380 (2019).
[11] Tahersima, M. H., Kojima, K., Koike-Akino, T., Jha, D., Wang, B., Lin, C., and Parsons, K., "Deep neural network inverse design of integrated nanophotonic devices," *arXiv preprint arXiv:1809.03555* (2018).
[12] Malkiel, I., Mrejen, M., Nagler, A., Arieli, U., Wolf, L., and Suchowski, H., "Plasmonic nanostructure design and characterization via deep learning," *Light: Science & Applications* **7**(1), 1–8 (2018).
[13] Peurifoy, J., Shen, Y., Jing, L., Yang, Y., Cano-Renteria, F., DeLacy, B. G., Joannopoulos, J. D., Tegmark, M., and Soljačić, M., "Nanophotonic particle simulation and inverse design using artificial neural networks," *Science advances* **4**(6), eaar4206 (2018).
[14] Ma, W., Cheng, F., and Liu, Y., "Deep-learning-enabled on-demand design of chiral metamaterials," *ACS nano* **12**(6), 6326–6334 (2018).
[15] Liu, Z., Zhu, D., Rodrigues, S. P., Lee, K.-T., and Cai, W., "Generative model for the inverse design of metasurfaces," *Nano letters* **18**(10), 6570–6576 (2018).
[16] Jiang, J. and Fan, J. A., "Simulator-based training of generative neural networks for the inverse design of metasurfaces," *Nanophotonics* (2019).
[17] An, S., Zheng, B., Shalaginov, M. Y., Tang, H., Li, H., Zhou, L., Ding, J., Agarwal, A. M., Rivero-Baleine, C., Kang, M., et al., "A freeform dielectric metasurface modeling approach based on deep neural networks," *arXiv preprint arXiv:2001.00121* (2020).
[18] Hodge, J. A., Mishra, K. V., and Zaghloul, A. I., "Rf metasurface array design using deep convolutional generative adversarial networks," in [*IEEE International Symposium on Phased Array Systems and Technology*], (2019).
[19] Ma, W., Cheng, F., Xu, Y., Wen, Q., and Liu, Y., "Probabilistic representation and inverse design of metamaterials based on a deep generative model with semi-supervised learning strategy," *Advanced Materials* **31**(35), 1901111 (2019).
[20] So, S. and Rho, J., "Designing nanophotonic structures using conditional deep convolutional generative adversarial networks," *Nanophotonics* **8**(7), 1255–1261 (2019).
[21] Jiang, J., Sell, D., Hoyer, S., Hickey, J., Yang, J., and Fan, J. A., "Free-form diffractive metagrating design based on generative adversarial networks," *ACS nano* **13**(8), 8872–8878 (2019).
[22] Kudyshev, Z. A., Kildishev, A. V., Shalaev, V. M., and Boltasseva, A., "Machine learning assisted global optimization of photonic devices," *arXiv preprint arXiv:2007.02205* (2020).
[23] Tao, Z., Zhang, J., You, J., Hao, H., Ouyang, H., Yan, Q., Du, S., Zhao, Z., Yang, Q., Zheng, X., et al., "Exploiting deep learning network in optical chirality tuning and manipulation of diffractive chiral metamaterials," *Nanophotonics* **1**(ahead-of-print) (2020).
[24] Asano, T. and Noda, S., "Iterative optimization of photonic crystal nanocavity designs by using deep neural networks," *Nanophotonics* **8**(12), 2243–2256 (2019).
[25] Christensen, T., Loh, C., Picek, S., Jakobović, D., Jing, L., Fisher, S., Ceperic, V., Joannopoulos, J. D., and Soljačić, M., "Predictive and generative machine learning models for photonic crystals," *Nanophotonics* **1**(ahead-of-print) (2020).

[26] Tang, Y., Kojima, K., Koike-Akino, T., Wang, Y., Wu, P., Tahersima, M., Jha, D., Parsons, K., and Qi, M., "Generative deep learning model for a multi-level nano-optic broadband power splitter," in [*Optical Fiber Communication Conference*], Th1A–1, Optical Society of America (2020).

[27] Blanchard-Dionne, A.-P. and Martin, O. J., "Successive training of a generative adversarial network for the design of an optical cloak," *arXiv preprint arXiv:2005.08832* (2020).

[28] Sajedian, I., Badloe, T., and Rho, J., "Finding the best design parameters for optical nanostructures using reinforcement learning," *arXiv preprint arXiv:1810.10964* (2018).

[29] Blanchard-Dionne, A.-P. and Martin, O. J., "Teaching optics to a machine learning network," *Optics Letters* **45**(10), 2922–2925 (2020).

[30] Blanchard-Dionne, A.-P. and Meunier, M., "Sensing with periodic nanohole arrays," *Advances in Optics and Photonics* **9**(4), 891–940 (2017).

[31] Blanchard-Dionne, A.-P. and Meunier, M., "Optical transmission theory for metal-insulator-metal periodic nanostructures," *Nanophotonics* **6**(1), 349–355 (2017).

[32] Gallinet, B., Butet, J., and Martin, O. J., "Numerical methods for nanophotonics: standard problems and future challenges," *Laser & Photonics Reviews* **9**(6), 577–603 (2015).

[33] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y., "Generative adversarial nets," in [*Advances in neural information processing systems*], 2672–2680 (2014).