Deep learning for inverse design of dielectric nanostructures with distinguishable RGB color signatures on dark-field microscopy

Juliette Jiménez Jaimes, Sofia Ponomareva, Peter Wiecha

Deep Learning for inverse design of dielectric nanostructures with distinguishable RGB color signatures on dark-field microscopy.

Juliette Jiménez Jaimes. Author[a], Sofia Ponomareva. Author[b], and Peter Wiecha . Author[c]

[a]Laboratoire Collisions Agrégats Réactivité (LCAR), 18 route de Narbonne, Toulouse, France
[b]Laboratoire d'analyse et d'architecture des systèmes (LAAS), 7, av. du colonel Roche, Toulouse, France
[c]Laboratoire d'analyse et d'architecture des systèmes (LAAS), 7, av. du colonel Roche, Toulouse, France

ABSTRACT

Silicon nanostructures have a rich optical response thanks to Mie-type optical resonances, that can be designed on-demand via their geometry. It is possible to encode bits of information in a nanostructure's geometry, and retrieve this information optically via the color observed in dark-field microscopy.[1] Furthermore, asymmetric structures can profit from the illuminating light polarization to facilitate information readout. Our ultimate goal is to accurately reverse engineer experimentally feasible silicon nanostructures for information encoding, such that they implement a set of ideally distinguishable colors for robust optical readout.

Deep learning is increasingly being used to solve inverse problems such as nano-photonic structure design. Neural networks for inverse design are mostly trained on simulated data, which is cheap to generate. But training neural networks on experimental data is a very interesting option, because it allows to include all experimental constraints into the model, which consequently learns to capture phenomena that may be hard to simulate.[2,3] Here, in order to learn an accurate model for the full experimental measurement setup, we trained a neural network with experimental darkfield color data from several thousand nanostructures.

Firstly, we built a forward network, taking as input the nanostructures' shapes from fabricated samples and predicting the dark-field color for both X and Y polarizations. We then successfully built an inverse tandem network, capable of designing structures with desired color responses. In order to create distinguishable color responses, another deep neural network was trained on the task to map all experimental colors in a regularized color latent space. Sampling equidistant points from this latent space then yields the most distinguishable, yet experimentally feasible colors.

The next future step will be to produce samples from the generated structures to test the network's accuracy. We would like to test how many bits of information we can encode using the darkfield color as readout.

Keywords: Neural networks for inverse design, neural networks trained on experimental data, darkfield colors for information encoding, inverse tandem network, silicon nanostructures, RGB response, color response.

1. INTRODUCTION

Sub-wavelength dielectric nanostructures are nowadays studied for their many properties that make them advantageous for high density information storage; such as fine-tuning of the optical response through the manipulation of the structure geometry[4] and their relatively easy and parallelizable manufacturing process. In a conceptually similar manner as for example in the CD-ROM, DVD or Blu-Ray disc, it is possible to encode multiple

Further author information: (Send correspondence to J.J.J.)
J.J.J.: E-mail: jimenez@irsamc.ups-tlse.fr
P.R.W.
S.P

bits of information in the nanostructure's geometry that itself is covering only a sub-diffraction small area.[5] Even though the structure cannot be resolved spatially from the far-field, the encoded information can be retrieved using spectrally resolved measurements, e.g., via the color observed in dark-field microscopy.[1] It has been demonstrated that artificial neural networks can be trained for accurate and robust information readout from the optical response of different 2D and 3D geometries of dielectric nanostructures.[1,6,7] Color encoding has the potential to enable fast, parallel and precise data extraction from a single measurement over a large area with many nanostructures; possibly enabling future industrialization of such technologies. However, a major challenge is to increase the number of digital bits of information that can be encoded in a single nanostructure, while remaining capable to retrieve this information from a fast color measurement. We try to address this open challenge using a deep learning based inverse design approach, which we introduce in the following.

2. METHODS AND RESULTS

2.1 Nanostructure Geometry

We want to model dielectric nanoparticles on top of a substrate, and observe the effect of changing their shape on the observed color on dark field images. We want to use nanostructures that do not have radial symmetry because we want these structures to produce different scattered spectra and different colors for X and Y polarized light. The structures are constructed with some cuboids with random lengths and widths (l_i, w_i), with l_i and $w_i \in [40\ nm - 150\ nm]$. The cuboids have random positions $(x_i, y_i) \in A$, with A being an area of 400 nm $\times$ 400 nm. These cuboids can intersect to form even larger and more complex figures, in which case we keep the union of the overlapping structures.

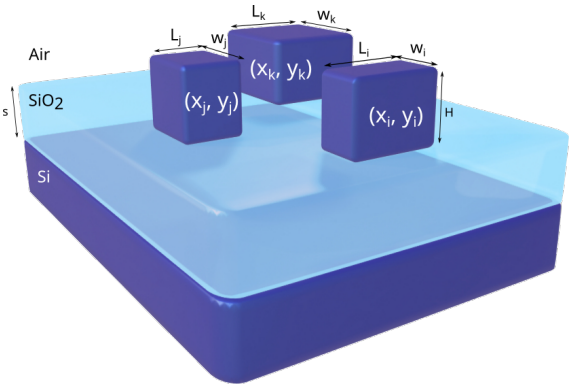

Figure 1. Representation of the Silicon nano cuboid structures.

2.2 Numerical Simulations

For the study and simulation of different dielectric nanoparticles on a substrate, and their interactions with light, Electromagnetic (EM) simulations were done using the following formalism of solving the inhomogeneous Helmholtz equation

$$\left(\mathbf{\Delta} + k^2\mathbf{1}\right)\vec{E}(\vec{r}) = \frac{-4\pi}{\varepsilon}\left(\mathbf{\Delta} + k^2\mathbf{1}\right)\chi_e\vec{E}(\vec{r}). \tag{1}$$

The solution to this equation, in presence of a nanostructure that occupies the volume V is given by the optical **Optical Lippmann-Schwinger** equation:[8]

$$\vec{E}(\vec{r},\omega) = \vec{E}_0(\vec{r},\omega) + \int_V \mathbf{G}^{EE}(\vec{r},\vec{r'},\omega)\chi_e(\vec{r'},\omega)\vec{E}(\vec{r'},\omega)\ d\vec{r'}\ , \tag{2}$$

$\vec{E}_0$ represents the illuminating field, $\vec{r'}$ is a position inside the nanostructure, V the nanostructure's volume; and $\mathbf{G^{EE}}$ the **Green's tensor**, describing the behavior of a point source in the given environment. Here we have a

layered substrate with semi-infinite Si and a $400nm$ SiO_2 (BOX) layer. This equation is numerically solved by discretizing the nanostructure volume into finite cells. For this purpose, we use the python toolkit `pyGDM`. This toolkit also allows us to calculate the back-scattering spectra of our system, which we then use to calculate the perceived color primaries (X, Y, Z) of the nanostructures when observed in dark field microscopy.[9,10] These color primaries are obtained using the CIE 1931 XYZ color matching functions, these are then converted to RGB values. Numerically, this treatment is done with the `colour` python toolkit.[11]

2.3 Experimental Sample Measurements

We fabricate an experimental sample with 8775 different silicon nanostructures, each with 5 micrometer distance to each other, on top of a Silicon-on-insulator substrate with a silicon layer height of $90nm$, with the same top-down electron-beam lithography process as used in prior works.[1] The sample was measured by dark field microscopy with a $NA = 0.45$ objective in back-scattering and a calibrated halogen lamp as white light source. Linear polarization filters are used for measuring X and Y polarized response. We take dark-field microscopy color images with fixed white balance of the sample and under the described, fixed experimental conditions to get each nanostructure RGB values.

However, we observed that many nanostructures produce a spot with multiple colors. Taking the mean color of all these shades often produces a very different color to what a human observer feels as perceived color when looking at the dark-field images. The experimental colors also generally don't match well with the colors from the back-scattering simulations. We attribute this mismatch between experiment and simulations to the non-negligible collection angle of our $NA = 0.45$ microscope objective, which leads to the different colors observed in the zoomed microscope images.

To resolve this wealth of contributing colors, we generate for each nanostructure a hierarchical color palette of the five most prominent colors observed in the dark field images, as seen in Fig. 2. This is done using k-means clustering and removing the "black" cluster. For this, we use the `scikit-learn` toolkit.[12] The hierarchical order allows us to characterize each structure by the one color that is predominant in the dark-field image. This predominant color information will be used to train our neural network.

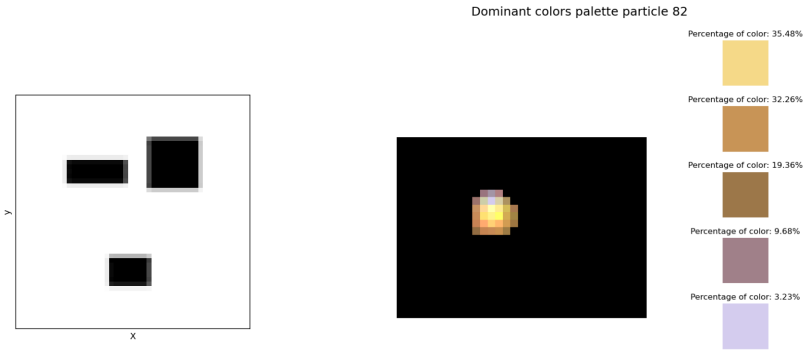

Figure 2. (a) Example of the mask image of a nanostructure. (b) Example of the color palette generated from the processed dark field image of this nanostructure.

While the simulations reproduce also similar palettes when analyzing angle-dependent color scattering, the agreement with experiment is rather qualitative, and not sufficient for an accurate identification of color-responses, as required for the target data storage application. Therefore, we try in a first step to train a data-based deep learning model for a quantitatively more accurate prediction of the experimentally observed colors.

2.4 Forward Network Architecture

We build a neural network model that analyses the nanostructure's mask image and predicts the observed color for both light polarizations.

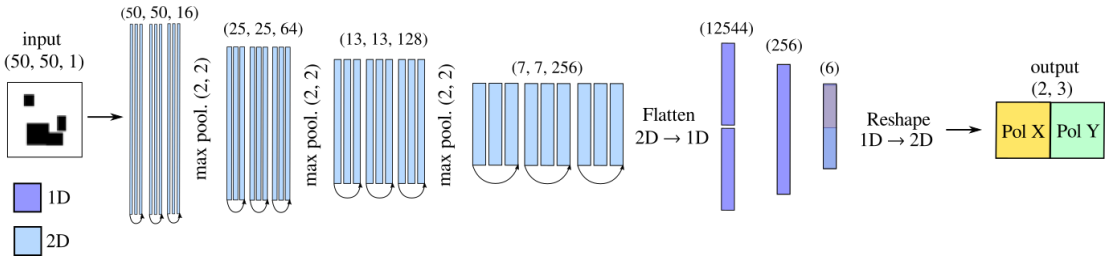

Figure 3. Diagram of the Forward Residual Network used for color prediction. The input of the ResNet is a $50px \times 50px$ image with values $\{0, 1\}$. The blue hidden layers, are 2D convolutional layers with residual blocks, that compress information from the mask image. The purple hidden layers are dense layers. The output of the network is an array containing RGB values for both light polarizations.

To avoid overfitting, we apply data augmentation by artificially incrementing the size of the training set by mirroring the mask images with respect to the X axis, The Y axis, and by randomly shifting the image structure by a couple of pixels. With this data augmentation, the model becomes more robust by learning that symmetry transformations of the mask image with respect to the incident field should leave the output color unchanged.

The created neural network is able to accurately predict the colors that we will observe for our specific experimental conditions. Note that the model can be easily modified to predict the main three hierarchical colors obtained for each nanostructure. This can be done by simply increasing the number of output channels to six (3 colors for two polarizations) and by adding the percentage of each color in the total response. Despite data augmentation, overfitting still occurs when we train directly on the experimental data. However, we manage to suppress overfitting by a simple trick: We train the same network on a significantly larger set of simulated colors first (20.000 samples), even though the simulated colors don't match well the experimental colors. The reasoning behind this approach is that the physics in the simulations is the same, and deviations with the measurements occur because we are not capable to reproduce the exact experimental conditions. We then fine-tune this simulations-based neural network with small learning rate on the experimental training data. In this way, we successfully avoid any overfitting and manage to obtain a smaller validation loss on the experimental test data, compared to experiment-only training.

For a first proof of principle, the following models are done only with one color input, in the future we plan to modify the workflow to include the richer color information.

2.5 Inverse Model

2.5.1 Tandem network

Our goal is to get a set of inverse designed nanostructures that encode a certain number of bits via a succinct and easy to measure color response, so that it is easy and robust to read out the information by comparing the colors. To this end, we have to develop an artificial Neural Network capable of predicting a nanostructure geometry that will yield a target Dark-field color.

Our nanostructures are parametrized as binary (black/white) top-view images, where black indicates silicon, and white indicates the absence of material. However, the vanilla tandem network (full model in Figure 4) struggles to generate "non-blurry" images, so we use a recently proposed trick:[2,13] We train a separate Wasserstein GAN with gradient penalty (WGAN-GP[14]) network on the generation of binary silicon nanostructure-images. The WGAN generates very accurate high contrast images with sharp edges. Its latent space is a learned mapping of the image-geometries to a continuous, parametrized description of the structures. The inverse design network (tandem architecture[15]) is subsequently trained to predict these z_{geo}, which then are converted by the pre-trained WGAN to a geometry-image. This image is then fed during the tandem-training into the pre-trained forward network for color-prediction and loss calculation.

A few examples of the inverse network's generated structures can be seen in Fig. 5.

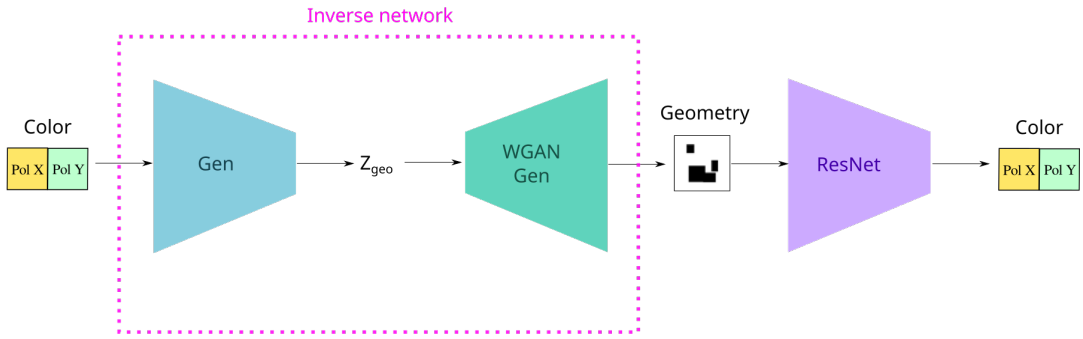

Figure 4. Diagram of the inverse network.

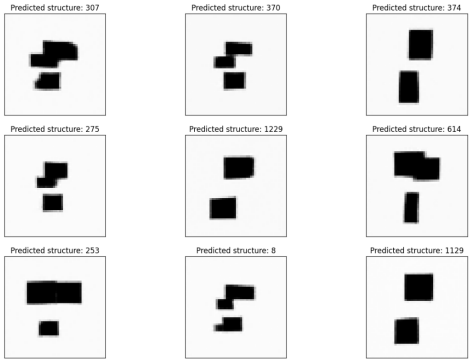

Figure 5. Example of generated geometries.

These top-view images look like the ones from the original data-set, as they are composed of blocks that can be joined to form more complex structures. The complete network manages to produce geometries whose colors are predicted to be very close to the original most predominant colors.

2.5.2 Variational autoencoder and latent space regularization

With the goal of finding the most "different" color combinations possible within the limits of the nano-geometry, a further autoencoder neural network is trained on mapping all experimental colors into a regularized color latent space (See Fig. 6, right). Due to the VAE regularization[16] in this latent space, similar colors would sit close to each other, and different colors would sit far away. The latent space regularization also allows us to get a compact and continuous dataset that allows interpolation between colors present in the original dataset (See Fig. 6, left).

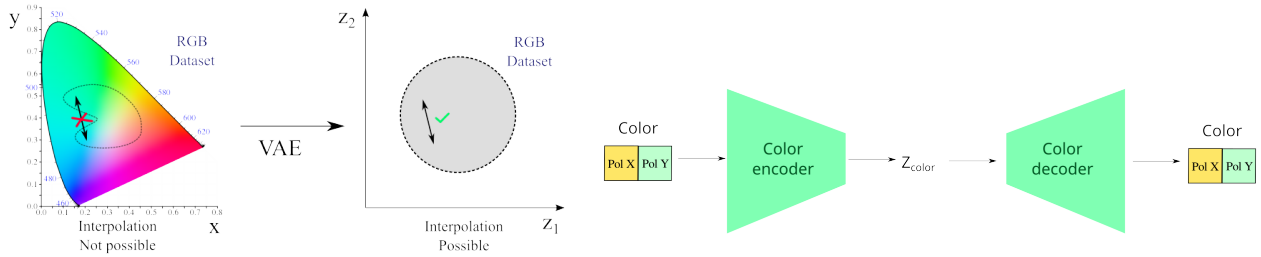

Figure 6. (a) Diagram of latent space regularization. (b) Diagram of Variation autoencoder.

An example of the latent space sampling is presented in Fig. 7 (left). The colors produced by the decoder of the Variational autoencoder, are mapped to show the network's prediction of the most predominant colors that

are feasible with our experimental setup (see Fig. 7, right). Finally, these feasible colors are fed to the tandem network (cf Fig. 4) to map the structures that would be needed to produce this distribution, which are depicted in the bottom of Fig 7.

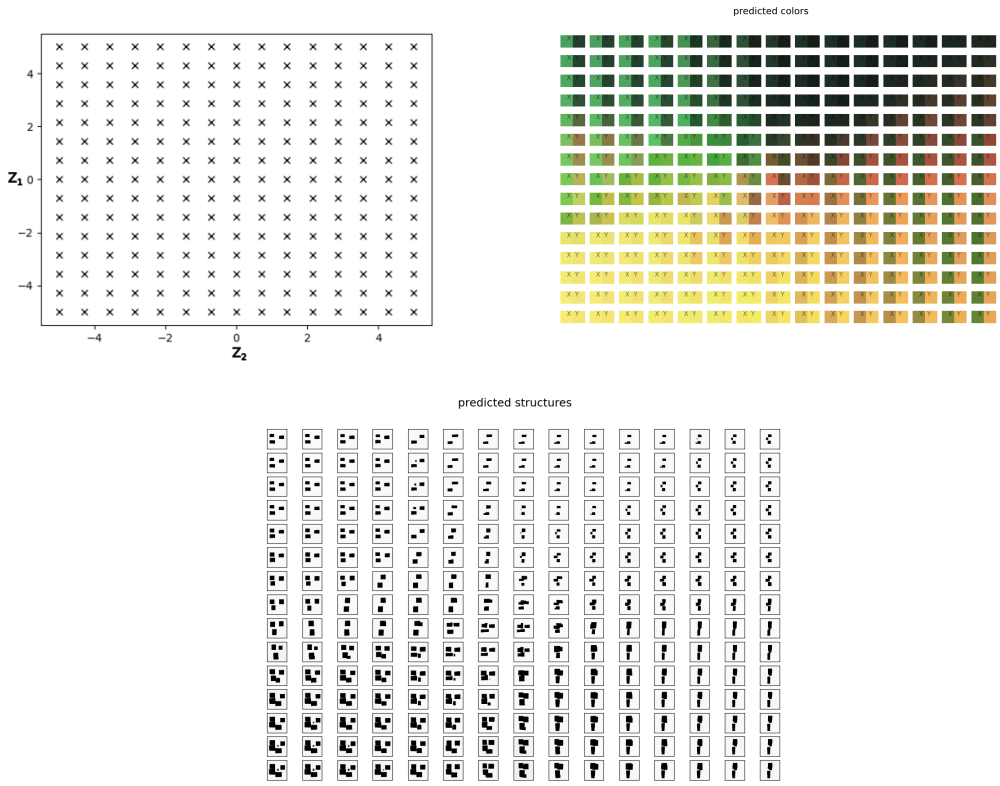

Figure 7. (a) Diagram of latent space sampling. (b) Predicted colors from this latent space sampling (c) Predicted nanostructure masks from this latent space sampling.

3. CONCLUSION

The aim of this work was to build a Deep Learning model for inverse design of dielectric nanostructures, targeting the color response of these structures to fabricate geometries with an optical signature that is as distinguishable as possible from other structures. An electromagnetic model was made, that captures the color tendencies produced by the nanostructures, and complex behaviors such as angular dependencies of the observed colors. We managed to construct and train a Neural Network on experimental data extracted from fabricated samples, to predict the dark-field observed color from dielectric nanostructures with different geometries. We also successfully built an Inverse Network capable of designing structures that will experimentally produce the desired optical responses for X and Y polarized light. Lastly, our goal is to use the here presented method to maximize the amount of bits that can be encoded and still retrieved in a single nanostructure.

REFERENCES

[1] P. R. Wiecha, A. Lecestre, J. D. Berke, and G. Larrieu, "Pushing the limits of optical information storage using deep learning," *Nature Nanotechnology* **14**, pp. 237–244, 1 2019.

[2] A. Khaireh-Walieh, D. Langevin, P. Bennet, O. Teytaud, A. Moreau, and P. R. Wiecha, "A newcomer's guide to deep learning for inverse design in nano-photonics," *Nanophotonics* **12**, pp. 4387–4414, 1 2023.

[3] P. R. Wiecha, A. Arbouet, C. Girard, and O. L. Muskens, "Deep learning in nano-photonics: inverse design and beyond," *Photon. Res.* **9**, pp. B182–B200, May 2021.

[4] A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk'yanchuk, "Optically resonant dielectric nanostructures," *Science* **354**, 11 2016.

[5] D. Psaltis and G. Burr, "Holographic data storage," *Computer* **31**(2), pp. 52–60, 1998.

[6] X. Li, Y. Cao, N. Tian, L. Fu, and M. Gu, "Multifocal optical nanoscopy for big data recording at 30 tb capacity and gigabits/second data rate," *Optica* **2**, pp. 567–570, Jun 2015.

[7] D. Yang, Z. Lei, L. Li, W. Shen, H. Li, C. Gui, and Y. Song, "High optical storage density using three-dimensional hybrid nanostructures based on machine learning," *Optics and Lasers in Engineering* **161**, p. 107347, 2023.

[8] C. Girard, "Near fields in nanostructures," *Reports on Progress in Physics* **68**, pp. 1883 – 1933, 2005.

[9] P. R. Wiecha, "pyGDM—A python toolkit for full-field electro-dynamical simulations and evolutionary optimization of nanostructures," *Computer Physics Communications* **233**, pp. 167–192, 12 2018.

[10] P. R. Wiecha, C. Majorel, A. Arbouet, A. Patoux, Y. Brûlé, G. C. D. Francs, and C. Girard, ""pyGDM" - new functionalities and major improvements to the python toolkit for nano-optics full-field simulations," *Computer Physics Communications* **270**, p. 108142, 1 2022.

[11] T. Mansencal, "Colour 0.4.2," *Zenodo* , 11 2022.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* **12**, pp. 2825–2830, 2011.

[13] Y. Augenstein, T. Repän, and C. Rockstuhl, "Neural Operator-Based Surrogate Solver for Free-Form Electromagnetic Inverse Design," *ASC photonics* **5, 10**, p. 1547–1557, 5 2023.

[14] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," *CoRR* **abs/1704.00028**, 2017.

[15] D. Liu, Y. Tan, E. Khoram, and ongfu Yu, "Training deep neural networks for the inverse design of nanophotonic structures," *ACS Photonics* **5, 4**, p. 1365–1369, 2 2018.

[16] D. P. Kingma and M. Welling, "An introduction to variational autoencoders," *CoRR* **abs/1906.02691**, 2019.