MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Deep Neural Network Inverse Modeling for Integrated Photonics

TaherSima, M.; Kojima, K.; Koike-Akino, T.; Jha, D.; Wang, B.; Lin, C.; Parsons, K.

TR2018-183 December 29, 2018

Abstract

We propose a deep neural network model that instantaneously predicts the optical response of nanopatterned silicon photonic power splitter topologies, and inversely approximates compact (2.6 x 2.6 um2) and efficient (above 92%) power splitters for target splitting ratios.

Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference (OFC/NFOEC)

Deep Neural Network Inverse Modeling for Integrated Photonics

Mohammad H. Tahersima, Keisuke Kojima*, Toshiaki Koike-Akino, Devesh Jha, Bingnan Wang, Chungwei Lin, Kieran Parsons

Mitsubishi Electric Research Labs. (MERL), 201 Broadway, Cambridge, MA 02139, USA.

**kojima @ merl.com*

Abstract: We propose a deep neural network model that instantaneously predicts the optical response of nanopatterned silicon photonic power splitter topologies, and inversely approximates compact ($2.6 \times 2.6\ \mu\text{m}^2$) and efficient (above 92%) power splitters for target splitting ratios.

OCIS codes: 250.5300, 160.3918, 100.4996.

1. Introduction

Subwavelength nanopatterned devices can be used to control incident electromagnetic fields into specific transmitted and reflected wavefronts [1–4]. However, optimization of such nanostructures, with a large design space is computationally costly. For example, computing the electromagnetic field profile via finite-difference time-domain (FDTD) methods may require long simulation time, several minutes to hours depending on the area of photonic device. To resolve the issue, we previously developed an artificial intelligence integrated optimization process using neural networks (NN) that can accelerate optimization by reducing the required number of numerical simulations [5, 6].

To design a photonic power splitter with arbitrary power ratios, photonic designers often begin with an overall structure based on intuition or analytical models, and fine tune the structure using a parameter sweep in numerical simulations [7, 8]. We demonstrate an alternative approach that uses deep learning methods to learn the large design space of a broadband integrated photonic power splitter in a compact deep neural network (DNN) model. To achieve this, we 1) investigate how to construct and train a DNN (8 layers), 2) apply an inverse design method to instantaneously obtain desirable performance, and 3) show that it is possible to design power splitters with multiple splitting ratios from the same trained DNN.

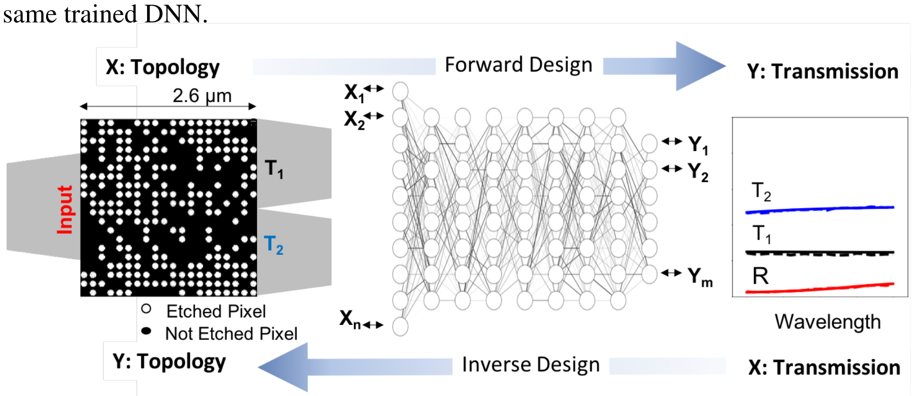

Fig. 1: By optimizing binary sequence of position of etch holes (white circles in the left figure) it is possible to manipulate light propagation towards either of ports. The DNN can take device topology as input and spectral response of the metadevice as label or vice versa.

2. Forward Prediction and Inverse Design

The goal of a nanostructured integrated photonics power splitter is to organize optical interaction events, such that the collective effect of the ensemble of scattering events guides the beam to a target port within target power intensity. To design the power ratio splitter using DNN we chose a simple three port structure on a standard fully etched SOI platform with a 220 nm-thick silicon layer. One input and two output 0.5μm waveguides are connected using an adiabatic taper to the 2.6μm wide square power splitter design region with a connection width of 1.3μm (Fig. 1). We use transverse electric (TE) mode as an input, and its conversion to transverse magnetic (TM) mode is minimal ($< 10^{-5}$).

To solve both forward and inverse design problems, we develop an eight layer deep and 100 neurons wide residual neural network (ResNet) [9]. The forward problem is approached as a regression problem while the inverse problem is solved as a classification problem, where we predict a binary vector representing the hole locations. A Gaussian log-likelihood function is used to train the DNN modeling the forward design problem. A Bernoulli log-likelihood classifier is used as the loss function for training the inverse problem. The Adam algorithm [10] is used for a training parameter optimization.

We then generated nearly 20,000 3D-FDTD simulation data. Some data are randomly generated in parallel, while others are generated sequentially through optimization algorithms such as Direct Binary Search. Each input data consists of 20×20 hole vectors (HV), each labeled by its spectral transmission response (SPEC) at port 1 and 2 and reflection from the input port. Each pixel is a circle with a radius of 45 nm and can have a binary state of 1 for etched ($n = n_{\text{Silicon}}$) and 0 for not etched ($n = n_{\text{Silica}}$). We split the data into 80 % for the training, and 20 % for testing. The training typically takes several hours using a desktop PC with a GPU board.

First we test the forward computation of the network to see prediction of spectral response of a topology data that the network is not trained on (Fig. 2 (a)). Interestingly, the network could predict transmission with above 99 % correlation coefficient (Fig. 2 (b)). This is a significant improvement from our previous results with only three layers [5]. Note that a conventional fully connected NN does not improve the performance beyond four layers, while the use of ResNet enabled us to increase the number of layers up to eight.

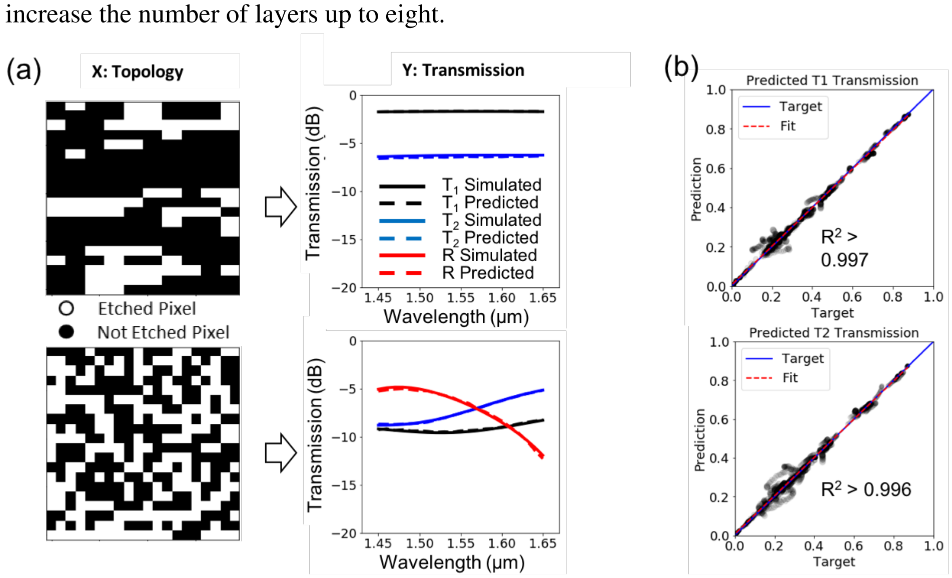

Fig. 2: (a) Spectral response of non-optimized nano-patterned power splitters, simulated by FDTD (solid lines) and predicted by DNN (dashed lines) at each output port for two device topologies, and (b) correlation between the target values (FDTD simulation) and the predicted values by DNN. Gray circle symbol size is proportional to gradient uncertainty.

Next, we test the inverse modeling on the same data as above by using SPECs as data and HVs as label and reversing and optimizing the inverse network. To test the generalization capabilities of the network, we investigate the network's inverse design performance on arbitrary and unfamiliar cases. To do this, we generate a reference table containing broadband constant transmission values for each port and use them as the input data batch for the Inverse Design DNN model. The predicted HVs can take any value from 0 to 1 from a Bernoulli distribution classifier. The classification converges to 0 or 1 as the loss reduces by increasing the number of training epochs. The predicted binary sequence is then fed back into the FTDT solver to confirm the validity of the response (Fig. 3). This inverse design process takes less than a second. The simulated total transmission efficiency is 96 %, 92 %, and 92 % for the splitting ratio of 1:1, 1:2, and 1:3, respectively. The spectral response is very flat for the whole wavelength range of 1450 – 1650 nm. The reflection is below 20 dB for most of the wavelength range.

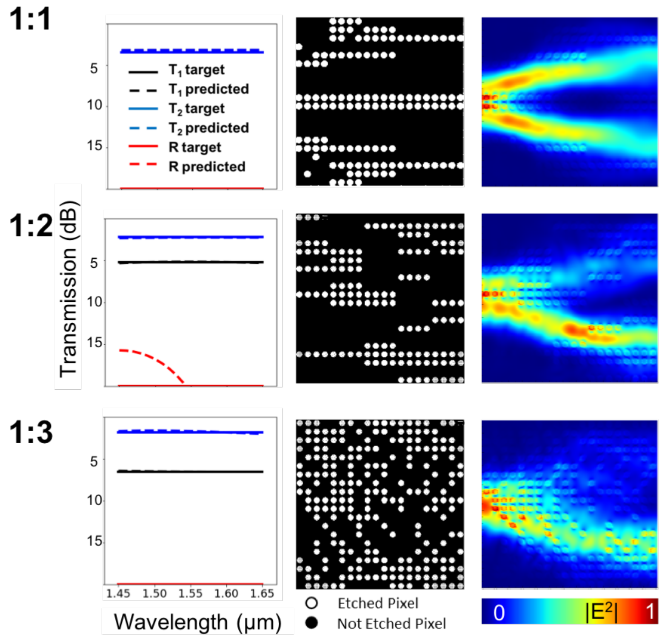

Fig. 3: Demonstration of inverse design using DNN with splitting ratios of 1:1, 1:2, and 1:3. Spectral response for transmission at port 1, 2, and reflection to the input port. FDTD-simulated values for predicted hole vectors on the right match the target values well.

3. Conclusion

DNNs can use device structure data (shape, depth, and refractive indices) as an input to predict the optical response of the nanostructure (forward network). In this case DNN can be used as a method for fast approximation of the optical response instead of using computationally heavy numerical methods. Another way to use DNNs, which is not available in the numerical method, is taking an optical response and providing the user with an approximate solution of nanostructure (inverse design). We demonstrate the application of DNNs in design of nanostuctured integrated photonic components. Although the design space for this problem is very large (2^{400} possible combinations), by training DNN with nearly 16,000 simulation data, we created a network that can approximate the spectral response of the an arbitrary hole vector within this design space. In addition, we could use the inverse network to design a nearly optimized power splitter topology for any user specific spectral responses, achieving > 92 % transmission efficiency and typically < 20 dB reflection for $1450 - 1650$ nm.

References

1. Ni, X., Wong, Z. J., Mrejen, M., Wang, Y., Zhang, X., "An ultrathin invisibility skin cloak for visible light," Science **349**, 1310-1314 (2015).
2. Arbabi, E., Arbabi, A., Kamali, S. M., Horie, Y., Faraon, A., "Multiwavelength polarization-insensitive lenses based on dielectric metasurfaces with meta-molecules," Optica **3**, 628-633 (2016).
3. Krasnok, A., Tymchenko, M., Alú, A., "Nonlinear metasurfaces: a paradigm shift in nonlinear optics," Materials Today **21**, 8-21 (2017).
4. Khorasaninejad, M., Chen, W. T., Devlin, R. C., Oh, J., Zhu, A. Y., Capasso, F. "Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging," Science, **352**, 1190-1194 (2016).
5. Kojima, K., Wang, B., Kamilov, U., Koike-akino, T., Parsons, K. "Acceleration of FDTD-based Inverse Design Using a Neural Network Approach," In *Integrated Photonics Research, Silicon and Nanophotonics* (Optical Society of America, 2017), paper ITu1A-4.
6. Teng, M., Kojima, K., Koike-Akino, T., Wang, B., Lin, C., Parsons, K. "Broadband SOI mode order converter based on topology optimization," In *Optical Fiber Communications Conference* (Optical Society of America, 2017), paper Th2A.8.
7. Tian, Y., Qiu, J., Yu, M., Huang, Z., Qiao, Y., Dong, Z., Wu, J.,"Broadband 1×3 Couplers With Variable Splitting Ratio Using Cascaded Step-Size MMI," IEEE Photonics Journal **10**, 1-8 (2018).
8. Xu, K., Liu, L., Wen, X., Sun, W., Zhang, N., Yi, N., Song, Q., "Integrated photonic power divider with arbitrary power ratios," Optics Letters **42**, 855-858 (2017).
9. He, K., Zhang, X., Ren, S., Sun, J., "Deep residual learning for image recognition," In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770-778 (2016).
10. Kingma, D. and Ba, J. A.: "A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).