Acceleration of FDTD-based Inverse Design Using a Neural Network Approach

Keisuke Kojima, Bingnan Wang, Ulugbek Kamilov, Toshiaki Koike-Akino, and Kieran Parsons
Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, MA 02139
kojima@merl.com

Abstract: Instead of using FDTD simulations for all the inverse design steps, we proposed to use neural network-based fitting to estimate the output of the FDTD simulations, and improve the design. We observed clear acceleration in the improvement of metric.

OCIS codes: (160.4236) Nanomaterials; 250.5300) Photonic integrated circuits;

1. Introduction

Nanophotonics and metamaterials are attracting more and more attention recently, due to increased availability of fabrication processes [1]. Subwavelength (SW) structures can suppress diffraction as long as the structure is smaller than half of the wavelength within the material. Many SW structures have been presented so far [2, 3]. Recently, Lu et al. reported an ultra-compact 1:2 coupler based on photonic crystal-like metamaterial structure [4]. The structure is relatively simple with 400 potential sites of voids (holes) and by determining whether a void will be created or not by using a 20x20 vector, the inverse design can be achieved by direct binary search (DBS). During the DBS process, they used 2D finite-difference time-domain (FDTD) simulations to save time, and later used 3D FDTD to fine tune the parameters. However, in various device simulation situations, 2D FDTD cannot always be a good indicator of 3D FDTD simulations. Therefore, it is desirable to use 3D FDTD simulations and still accelerate the inverse design process. We propose a new inverse design process using a fitting function. In our case, we use a simple two-layer neural network and some training data to estimate the output of the FDTD simulations. Therefore, we can expected accelerated design without conducting FDTD simulations for all the cases.

2. Simulation Procedures

2.1. *Device Structure*

We use the same 1:2 coupler/splitter as in [4], for the evaluation of our proposed method. We start with a 220 nm-thick silicon layer deposit on a SiO2 layer. The device region is $2.6\mu m \times 2.6\mu m$, and is discretized into 20×20 pixels. We take advantage of the symmetric structure, so the simulation time is about half, and the search space is much smaller. The input and output waveguides are 0.5 μm-wide, and the spacing between the two output waveguides is 1 μm. These pixels are represented by an air hole (void) with a radius of $0.47\mu m$ and an etching depth of 140 nm. Since we only treat symmetric cases, 10×20 binary matrix represent the device structure. One 3D FDTD simulation takes about 30 seconds on a high-end PC using Lumerical FDTD, depending on the mesh size. Our metric is the transmittance (1 being perfect transmittance). Excess loss is a loss beyond the 3 dB intrinsic loss.

2.2. *Conventional method*

For the baseline conventional cases, we use the direct-binary-search (DBS) algorithm. From the first pixel, we toggle the state of the pixels, and if the result is better, we retain the results. If the result is worse, then we keep the old state. We scan over the pixel locations multiple times to ensure conveergence.

2.3. *Proposed method*

In our proposed method we use a simple two-layer neural network model, as shown in Fig. 1 using the fitting function of Matlab Neural Network Toolbox. Fig. 2 shows the testing result where 700 input data is used for training, while testing was done with 150 input data. The the unseen test data, correlation between the target (FDTD simulation data) and the NN output is as high as 0.79448 in this case, indicating good predictability. We use a hybrid NN-DBS system, where the output of the neural network is used as the estimate and perform the DBS over the whole 200 pixels. Since training takes about 5 seconds (depending on the number of data set), and the fitting is almost instantaneous, the time for this hybrid NN-DBS process is negligible compared to a FDTD simulation. After this hybrid NN-DBS process, if the predicted result is better than the previous FDTD result, we use an FDTD simulation to verify the results. If the results is better than the previous FDTD simulation results, we retain the results. If the result is worse, then we retain the previous FDTD results. If the predicted result from the hybrid NN-DBS is worse than the previous FDTD result, then we revert to a conventional FDTD-based DBS for this step. As we accumulate the FDTD results, we used online learning, i.e., the newly acquired data are included in the training data set. In order to increase weight for the new data, all the data after the initial 700 data were included in the training data 7 times.

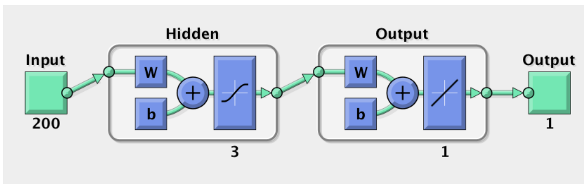

Fig. 1: Two-layer neural network structure. In this case, three neurons were used for the hidden layer.

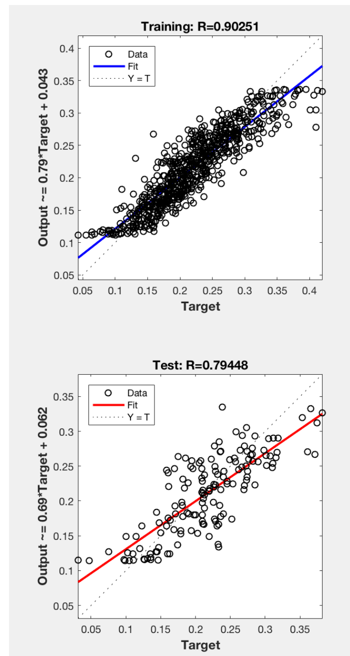

Fig. 2: Example of the NN fitting. NN is trained using 700 input data, while testing was done with 150 input data.

3. Simulation Results

We compared the conventional and the proposed methods using three initial conditions. For the proposed method, we used 700 randomly-run training data. The justification of the data are two-folds. First, these training data can be independently run, therefore they can be prepared simultaneously and quickly on a cluster. This is in contrast to the DBS case where all the simulations have to be done sequentially. Second, we will be running multiple FDTD simulations in any way during the course of optimization, so the data can be reused.

Fig. 3 shows the excess loss as a function of the number of FDTD runs. As you can see, there is a great acceleration in the convergence of excess loss in the case of our proposed method especially in the early stage of the convergence. In fact, in our 700 training data, the best excess loss is 3.8 dB, so we are not unfairly using the prior knowledge directly to reduce the excess loss. The acceleration is purely from the fitting/prediction power of the neural network method.

Fig. 4 shows optimized structure using the proposed method, with the initial condition 3. There is clearly a structure in the final result. The simulated excess loss spectrum is shown in Fig. 5. For the range of 1500 nm-1600 nm, the excess loss below 0.4 dB is expected, which is comparable to those reported earlier [4].

4. Conclusion

We propose a novel method of using neural network as a predictor of the FDTD simulation. The inverse design process can be significantly accelerated due to this proposed method.

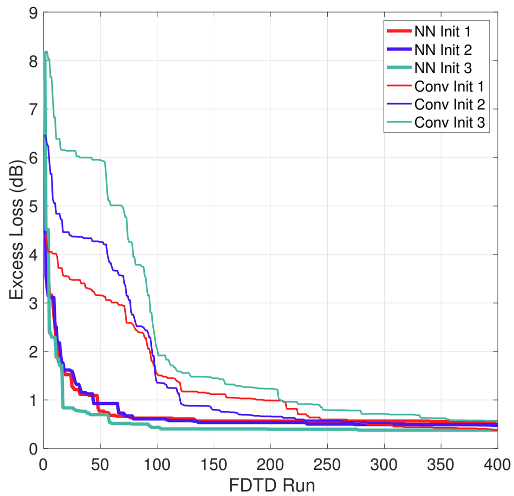

Fig. 3: Convergence of excess loss with a function of the number of FDTD simulation runs, with three different initial conditions. NN stands for neural network, while Conv stands for the conventional direct binary search

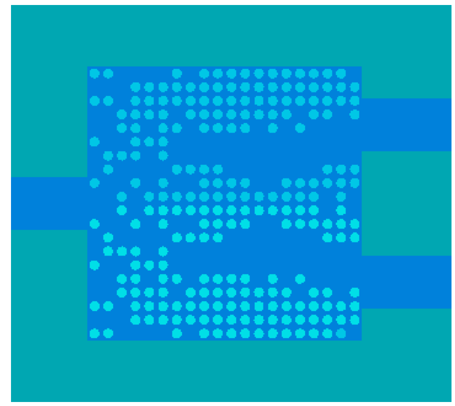

Fig. 4: Optimized structure using the proposed neural network method starting from the initial condition 3.

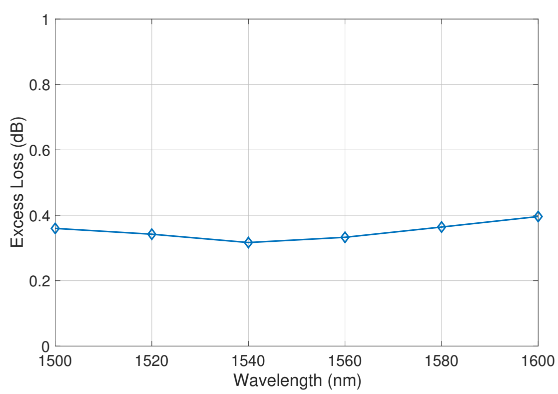

Fig. 5: Simulated excess loss vs. wavelength of the device optimized by the proposed method.

References

1. P. Cheben et al, "Subwavelength Index Engineered Waveguides and Devices," Proc. OFC, Tu3K.2, Los Angeles (2017).
2. A. Y. Piggott, et al., "Inverse design and demonstration of a compact and broadband on-chip wavelength demultiplexer," Nature Photonics 9, 374 (2015).
3. B. Shen, et al., "An integrated-nanophotonics polarization beamsplitter with $2.4 \times 2.4\ \mu m^2$ footprint," Nature Photonics 9, 378 (2015).
4. L. Lu, et al., "An ultra-compact colorless 50:50 coupler based on PhC-like metamaterial structure," Proc. OFC, Tu3E.5, Anaheim (2016).