B-2-05 Extended Abstracts of the 2019 International Conference on Solid State Devices and Materials, Nagoya, 2019, pp75-76

Computational design of high-efficiency grating coupler based on deep learning

Yuto Miyatake, Naoki Sekine, Kasidit Toprasertpong, Shinichi Takagi and Mitsuru Takenaka

Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8656, Japan
Phone : +81-3-5841-6733, E-mail : miyatake@mosfet.t.u-tokyo.ac.jp

Abstract

High-efficiency grating coupler on a standard 220 nm SOI wafer without back reflector is designed based on deep learning. The neural network enables us to explore the wide design-parameter space of the grating coupler. As a result, we achieve the peak coupling efficiency of 73.6%.

1. Introduction

Silicon (Si) photonics is one of the most interesting platforms for the next-generation information processing and telecommunication. A Si-on-insulator (SOI) structure offers two notable advantages: strong optical confinement owing to the high refractive index contrast between Si (n=3.5) and SiO_2 (n=1.45) and complementary-metal-oxide-semiconductor (CMOS) compatibility. These features allow ultracompact integration of optical components, making it possible to realize large-scale photonic integrated circuits (PICs) [1]. As a PIC comprises many optical components, their performances directly affect the whole performance of a PIC. Design of optical components, however, has inevitably been accompanied by a restriction of the number of design parameters based on designer's consideration and/or intuition because the design-parameter space for optical components is too large for brute-force search using time-consuming FDTD simulation. In particular, the design of grating couplers encounters this restriction. Although the coupling efficiency of grating couplers has been improved gradually [2][3], there is still room for exploring its design-parameter space more deeply.

In this paper, we examine to use deep learning for designing an efficient grating coupler on a standard 220 nm SOI wafer without a back reflector. The design optimization can be accelerated by using neural networks (NN), which have recently been used to approximate many physics simulations [4][5]. As a result, we find high-efficiency grating coupler with a peak coupling efficiency of 73.6%.

2. Numerical model and results

In this work, we optimize the coupling efficiency of a grating coupler between a single-mode optical fiber and the fundamental TE mode of a Si waveguide. We assume a standard SOI wafer with a 220-nm-thick Si layer and a 2-µm-thick buried oxide (BOX) layer as shown in Fig. 1. The thickness of the top SiO_2 cladding is 720 nm. We assume that there is no back reflector at the bottom of the BOX layer. The incident angle of the single-mode optical fiber is fixed to be 10°. We define the grating as 24 trench widths and 24 tooth widths. Therefore, these 48 design parameters are optimized to maximize the coupling efficiency. The etching depth of the grating is fixed to be 105 nm in this simulation.

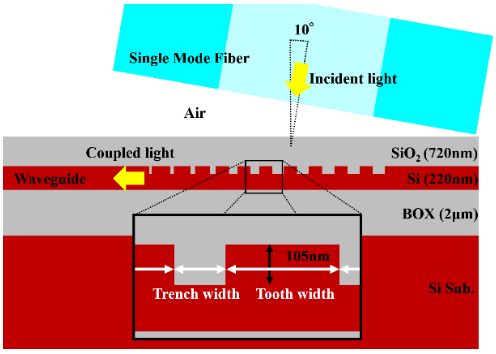

Fig. 1 Cross-sectional schematic layout of a grating coupler on a standard SOI wafer.

In the first step of the design optimization, we train a NN on a dataset to predict coupling efficiency from the design parameters as shown in Fig. 2. Once trained, the NN can calculate the gradient of the output with respect to the input parameters analytically instead of numerically. Then, in the second step, we optimize the design parameters using the trained NN by a gradient descent method.

In the first step, we use 2D-FDTD simulations to generate 5,000 training instances, each of which comprises a design parameter and its coupling efficiency spectrum. We obtain randomly-fluctuated design parameters by adding random displacements to the seed structure that is designed according to [3]. We use 3,500 data for training, while other 1,000 and 500 data are left as a validation dataset and a test dataset, respectively. We use a fully connected network, with three layers and 48 units per layer since the input data are 48-dimentional vectors instead of tensors. The outputs are spectra sampled at 100 points between 1.5 µm and 1.6 µm wavelengths. The activation function is ReLU and the dropout ratio is 0.01. These hyperparameters are determined by shuffle-split-cross validations.

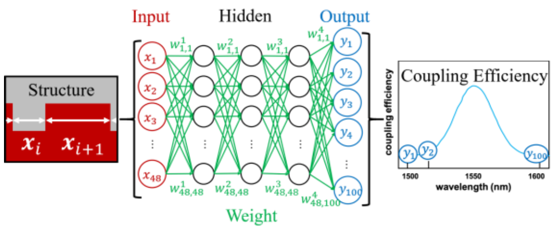

Fig. 2 The neural network architecture has the trench widths and tooth widths as its inputs and the coupling efficiencies at different wavelengths as its outputs.

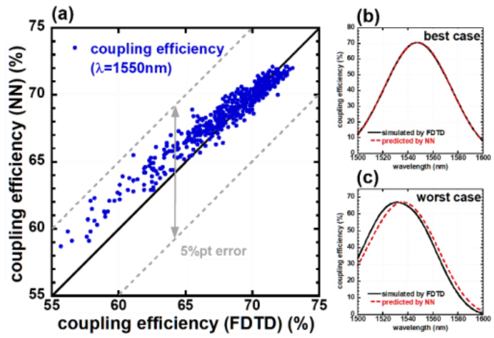

Fig. 3 (a) Predicted and true coupling efficiency values at a wavelength of 1550 nm for the test dataset. The predicted and true coupling efficiency spectra of (b) the best case and (c) the worst case.

The relationship between predicted and true coupling efficiency at a wavelength of 1550 nm is shown in Fig. 3. We observe that the error is lower than 5%pt for all the data included in the test dataset and predicted coupling efficiency spectrum is very close to the true one even in the worst case as shown in Fig. 3(c).

In the second step, we optimize the design parameters by a gradient method, where the gradient of a loss function with respect to the design parameters is calculated using the trained NN as shown in Fig. 4. The loss function $Loss$ is given by

$$Loss = -CE + p \cdot \|\boldsymbol{x}_t - \boldsymbol{x}_{\text{intial}}\| \quad (1)$$

where CE is the coupling efficiency at a wavelength of 1550 nm and p is a penalty parameter. The second term of the loss function is the product of the penalty parameter and the Euclidean distance between the initial structure and the current structure. We optimize the penalty parameter to suppress the large displacement from the initial seed structure. The number of learning iterations is set to 10,000. We selecte Adam (adaptive moment estimation) amang several gradient methods because Adam is known to show the fastest convergence in many problems [6]. Note that we can use a gradient-based method because the trained NN can quickly calculate the gradient analytically instead of numerically; otherwise, we can only use a derivative-free method which requires more function counts than a gradient-based method.

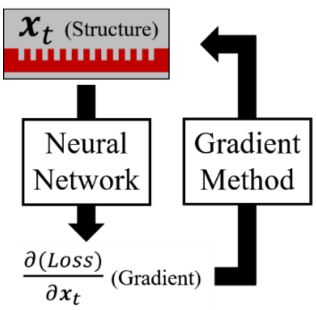

Fig. 4 Gradient-based optimization using the trained neural network.

Figure 5 shows the convergence of the coupling efficiencies calculated by FDTD and the NN. The calculation by FDTD was performed after the optimization for validation. The values calculated by the NN slightly deviate from that calculated by FDTD as they converge. This is due to the inaccuracy of the NN's calculation. Nevertheless, we observe relatively smooth convergence during 15,000 iterations because the NN can properly evaluate the gradient of the loss function. Figure 6 shows the coupling efficiency spectra of the initial and the optimized structure. We achieve the maximum coupling efficiency of 73.6% at a wavelength of 1550 nm, which is one of the best efficiencies among grating couplers on a standard SOI wafer with no back reflector. Thus, we successfully explore a wider design-parameter space of a grating coupler than ever before using deep learning.

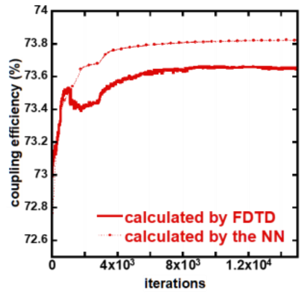

Fig. 5 Convergence of the coupling efficiencies calculated by FDTD and NN.

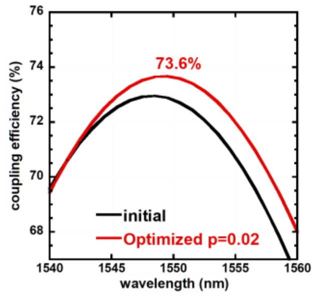

Fig. 6 Coupling efficiency spectra of the initial and optimized structure.

3. Conclusions

We have designed high-efficiency grating coupler on a standard 220 nm SOI wafer without back reflector based on deep learning and demonstrated the maximum coupling efficiency of 73.6%. The examined design method is also promising for designing other optical components.

Acknowledgements

This work was partly commissioned by the New Energy and Industrial Technology Development Organization (NEDO).

References

[1] B. Jalali and S. Fathpour, J. Lightw. Technol. **24** (2006) 4600-4615

[2] D. Taillaert et al., Japan. J. Appl. Phys. **45** (2006) 6071-6077

[3] R. Marchetti et al., Sci. Rep. **7** (2017) 16670.

[4] J. Peurifoy et al., Sci. Adv. **4** (2018) eer4206

[5] T. Asano and S. Noda, Opt. Express **26** (2018) 32704-32716

[6] D. Kingma and J. Ba, arXiv:1412.6980v8(2015)