Realization of high-speed 3D waveguide analysis model via transfer learning based on 2D-FDTD simulation

Gai Ichisawa
Institute of Science Tokyo, Japan
ichisawa.g.b63a@m.isct.ac.jp

Sho Okada
National Institute of Information and Communications Technology, Japan
okada.s.ah@nict.go.jp

Tomohiro Amemiya
Institute of Science Tokyo, Japan
amemiya.t.e262@m.isct.ac.jp

***Abstract*—In this study, we realized an optical device analysis program that achieves results equivalent to those of 3D-FDTD with a small amount of computation time and computational resources using transfer learning, and actually applied it to a 1x2 MMI coupler to discuss its usefulness. The proposed model has an R^2 score of 0.901 and an analysis time of about 47.4 µs per calculation, which is sufficiently fast compared to the 3D-FDTD.**

Keywords—FDTD, DNN, Transfer Learning, Side Tuning

I. INTRODUCTION

Optical circuit technology, including silicon photonics, is becoming indispensable for future optoelectronic convergence because it enables high-speed, large-capacity, and low-power data transmission. In this context, the finite difference time domain (FDTD) method is becoming more important for designing various devices in optical circuits. However, a bottleneck in the optimal design of devices using 3D-FDTD is that it requires a large amount of time for analysis (or a large amount of computational resources), even for simple structures.

Against this background, in this study, we realized an optical device analysis program that achieves results equivalent to those of 3D-FDTD with a small amount of computation time and computational resources using transfer learning, and actually applied it to a 1x2 MMI coupler to discuss its usefulness.

II. METHODS OF REALIZING ANALYSIS PROGRAM

When realizing an optical device analysis program that uses deep learning to obtain results equivalent to those of 3D-FDTD, it is necessary to collect a large amount of training data using 3D-FDTD for the target optical device in advance. However, it is clear that this is not realistic in terms of computation time and computational resources.

Therefore, in this study, we introduced a method in which a neural network is roughly trained using a large amount of analysis data with 2D-FDTD, which has a low computational load, and then the network is optimized by transfer learning using a small amount of analysis data with 3D-FDTD. This method is versatile regardless of the type of optical device, and in this study, its effectiveness was verified by creating an analysis program for a 1x2 multi-mode interferometer (MMI) coupler.

III. DATA PREPARATION

As mentioned above, a deep neural network (DNN) analytical model for a 1x2 MMI coupler was created in this study. As shown in Fig. 1, four structural parameters were set as input variables: length L, width W, and output port positions P_{o1} and P_{o2} (-1 for the end of the device, 0 for center, and 1 for the other end of the device), and the output variable was the light intensity from each port.

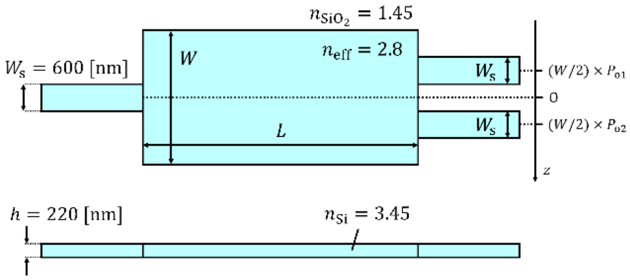

Fig. 1 Structure of MMI applying the proposed analysis method using transfer learning.

Table 1 Structural parameter distribution settings

Parameter	Average	Standard deviation	Range
L	15.5 μm	4.00 μm	[3.50,27.5] μm
W	3.00 μm	0.33 μm	[2.00,4.00] μm
P_{o1}, P_{o2}	0	0.33	$[-1.0, 1.0]$

In this study, we generated 10000 sets of 2D-FDTD data and 400 sets of 3D-FDTD data using *Synopsys RSoft Photonic Device Tools* while randomly varying the input variables to stay within ±3σ according to the normal distribution in Table 1 (including moderate amounts of structural parameters for devices with high output intensity).

IV. NETWORK DESIGN FOR 2D-FDTD

First, a neural network was trained to predict the intensity of the two output ports of a 1x2 MMI coupler using 1000 sets of 2D-FDTD data. For training, we used 5-fold cross-validation, with the average of the maximum R^2 scores as the evaluation index for the network. Models that were determined to be overlearning by looking at the loss function were excluded from the evaluation.

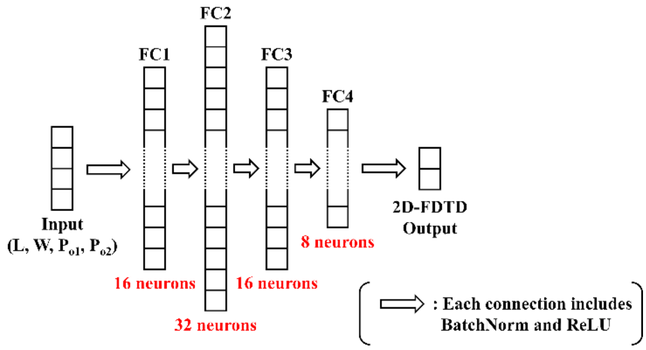

Fig. 2 Network configuration trained with analyzed data from 2D-FDTD.

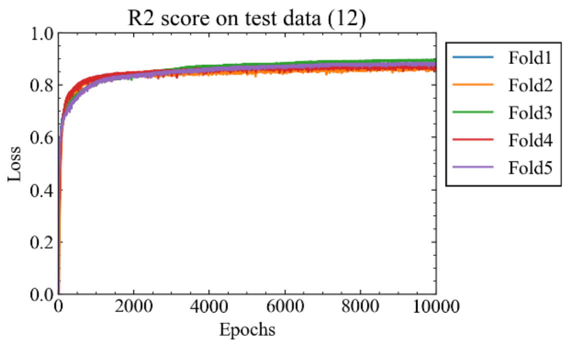

Fig. 3 Score transition of NN shown in Fig. 2.

The optimized network and score transition are shown in Figs. 2 and 3. The optimized DNN was fully connected with 16, 32, 16, and 8 nodes in the hidden layer, where the R2 score was 0.936

V. Network Design for 3D-FDTD

Next, based on the network trained with the analysis data from the 2D-FDTD, the network was optimized by transfer learning [1-3] using a small amount of analysis data from the 3D-FDTD. The final DNN analysis model for the 1x2 MMI coupler is shown in Fig. 4. First, the four structural parameters are input to the DNN trained in section IV (Fig. 2) to obtain two feature values. Next, these two feature values and the original four inputs are combined to form a 6-dimensional vector. This vector is input to a small network (2 layers, 8 nodes per layer) to finally obtain the light intensity from each port. After training the small network, the R^2 score was 0.901, as shown in Fig. 5.

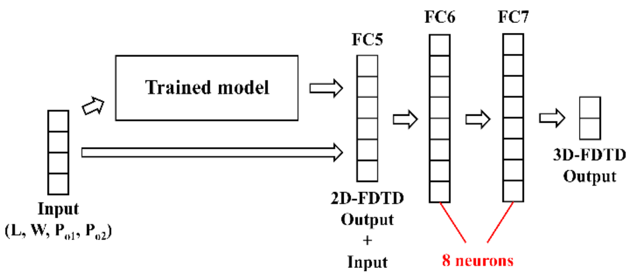

Fig. 4 Transfer learning network configuration trained with analyzed data from 3D-FDTD.

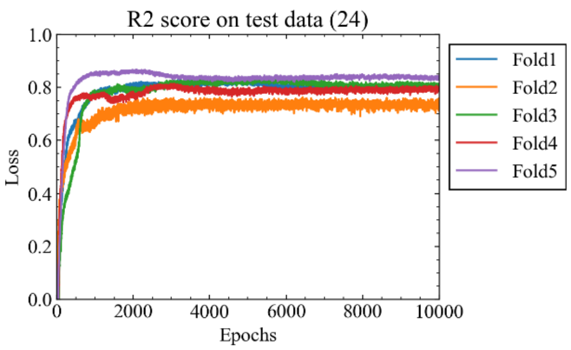

Fig. 5 Score transition of NN shown in Fig. 4.

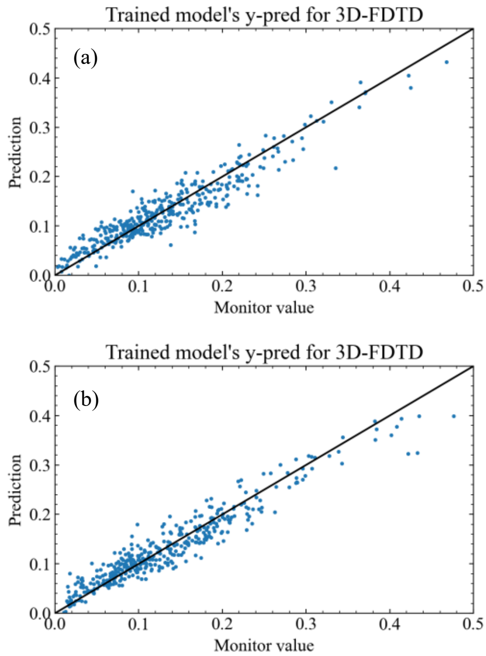

Fig. 6 A plot of the actual 3D-FDTD analysis results and the DNN prediction results obtained in Section V. (a) Output port 1. (b) Output port 2.

Table 2 Average calculation time for each method

Method	Software	Processor	Time
3D-FDTD	Synopsys RSoft Photonic Device Tools	Intel Xeon Gold 6238R	3 h
DNN	Python (PyTorch)	NVIDIA GeForce RTX 3070	47.4 µs

VI. DNN Performance

A plot of the actual 3D-FDTD analysis results and the DNN prediction results obtained in section V is shown in Fig. 6. 10^5 runs of DNN predictions were performed, and all calculations were completed in 4.74 s, for an average analysis time of 47.4 µs per run. It can be concluded that by using transfer learning, analysis results equivalent to 3D-FDTD could be obtained with less computation time and computational resources.

This method is considered to be versatile because it can be used not only for MMI but also for other optical devices.

Acknowledgment

This work is supported by JSPS KAKENHI (25H00734), JST CREST (JPMJCR24R1), JST ASPIRE (23837157).

References

[1] F. Zhuang, *et al.*, "A Comprehensive Survey on Transfer Learning," in *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43-76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.

[2] M. Iman, H. R. Arabnia, and K. Rasheed, "A review of deep transfer learning and recent advancements," *Technologies*, vol. 11, no. 2, p. 40, 2023, doi: 10.3390/technologies11020040.

[3] J. O. Zhang, A. Sax, J. Malik, L. Guibas, and A. Zamir, "Side-tuning: A baseline for network adaptation via additive side networks," *Computer Vision – ECCV 2020*, pp. 698–714, 2020, doi: 10.1007/978-3-030-58580-8_41.