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

Nano-Optic Broadband Power Splitter Design via Cycle-Consistent Adversarial Deep Learning

Tang, Yingheng; Kojima, Keisuke; Koike-Akino, Toshiaki; Wang, Ye; Jha, Devesh; Parsons, Kieran; Qi, Minghao

TR2021-045 May 11, 2021

Abstract

A novel generative deep learning model with a cycle-consistent adversarial network is introduced for optimizing 550 nm broad bandwidth (1250 nm to 1800 nm) power splitters with arbitrary target splitting ratios.

Conference on Lasers and Electro-Optics (CLEO)

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

Nano-Optic Broadband Power Splitter Design via Cycle-Consistent Adversarial Deep Learning

Yingheng Tang[1,2], Keisuke Kojima[1,*], Toshiaki Koike-Akino[1], Ye Wang[1], Devesh K. Jha[1], Kieran Parsons[1] and Minghao Qi[2]

[1]Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA.
[2]School of Electrical and Computer Engineering and Birck Nanotechnology Center, Purdue University, West Lafayette, IN 47907, USA.
[*] kojima@merl.com

Abstract: A novel generative deep learning model with a cycle-consistent adversarial network is introduced for optimizing 550 nm broad bandwidth (1250 nm to 1800 nm) power splitters with arbitrary target splitting ratios.

OCIS codes: (130.3120) Integrated optics devices; (160.1245) Artificially engineered materials

1. Introduction

Photonic device design often involves a high-dimensional search space and heavy simulation costs. This has prompted widespread interest in deep neural networks (DNNs) for design automation. However, DNNs generally require a large amount of training data, which can be quite expensive for silicon photonics problems. Our earlier work [1] employing an adversarial conditional variational autoencoder (A-CVAE) model showed the capability to generate the multilevel nanostructured splitters with arbitrary splitting ratios. The generative model combined with variable sized holes enable the splitters to achieve decent performance with an overall transmission of near 90%. In this paper, we introduce an improved DNN model which incorporates 'cycle consistency' [2–4]. The proposed model can improve the device performance in both total transmission and splitting ratios.

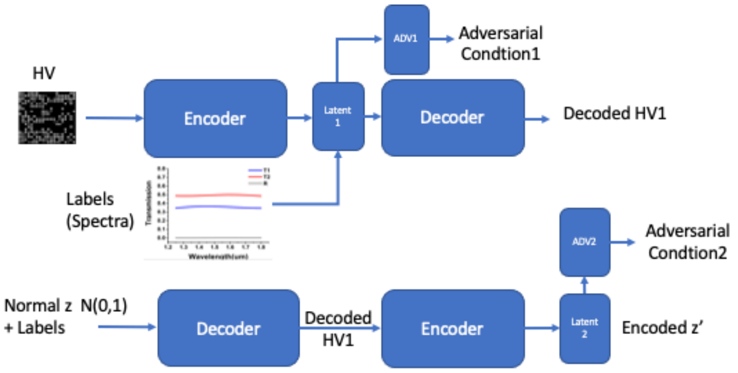

Fig. 1: The generative deep learning model structure of cycle-consistent A-CVAE.

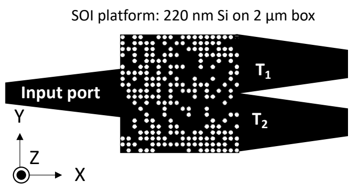

Fig. 2: Nanophotonic power splitter.

2. A-CVAE model with cycle consistency

The concept of A-CVAE model with cycle consistency [2, 4] is that in addition to training the encoder-decoder pair for forward reconstruction, the reversed-order decoder-encoder pipeline is also trained for latent space consistency (respective DNN modules share the same weights) as shown in Fig. 1. In this case, we are not only comparing the decoded Hole Vectors (HVs) but also the latent variables. Under the cycle-consistency idea, we explicitly train the encoder to reduce leakage of information associated with specified factors of latent variation. By doing this, we expect to further improve the generation performance of the original A-CVAE model. The loss function is shown as follows:

$$\mathrm{Loss} = a \cdot \sum_{i=1}^{n}(x_i - x_i')^2 + b \cdot \frac{1}{2}\sum_{j=1}^{J}\left[\mu_{zj}^2 + \sigma_{zj}^2 - \log(\sigma_{zj}^2) - 1\right]$$
$$+ c \cdot \sum_{i=1}^{n}(z_i - z_i')^2 + \alpha \cdot \sum_{i=1}^{n}(s_i - \bar{s}_i)^2 + \beta \cdot \sum_{i=1}^{n}(s_i - s_i')^2, \tag{1}$$

where the first term is the mean-square error (MSE) loss between the decoded HV1 and actual HV, the second term is the Kullback–Leibler divergence of the CVAE, the third term is the newly added MSE between the two latent spaces z and z'. The last two terms are the loss functions of the two adversarial blocks. The splitters are designed on a 220 nm thick silicon-on-insulator (SOI) multimode interference (MMI) footprint with a dimension of $2.25 \times 2.25\ \mu\mathrm{m}^2$ having 20×20 etched holes (Fig.2).

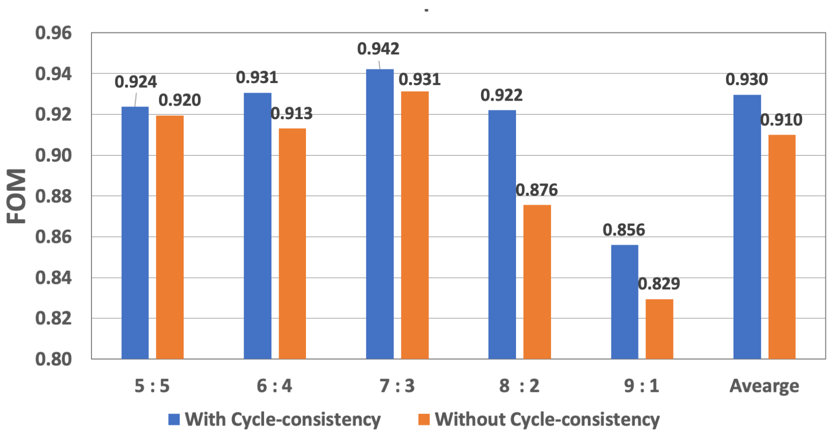

Fig. 3: FOM comparison for two generative models: A-CVAE model with cycle consistency (blue) and without cycle consistency (orange).

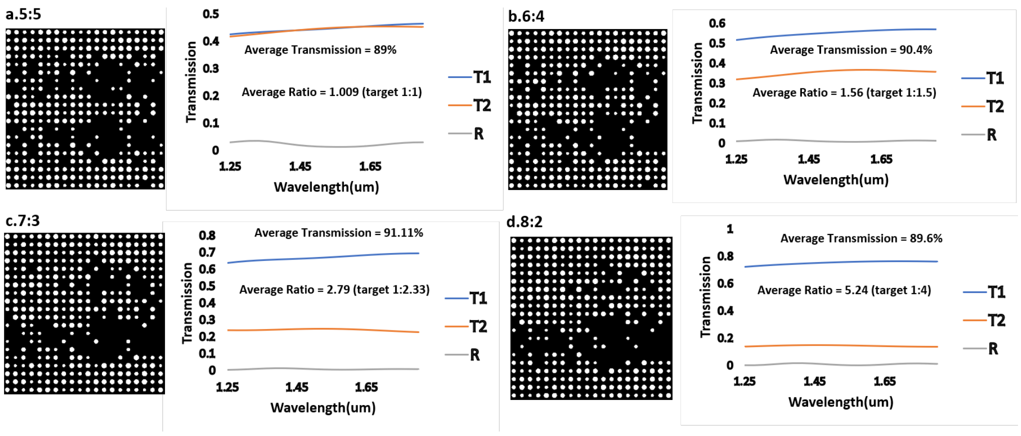

Fig. 4: FDTD results (transmission and reflection) of some generated patterns via the A-CVAE model with cycle consistency.

In order to better measure the performance of the generated devices, we define a figure of merit (FOM) to evaluate the performance of the designed splitters:

$$\mathrm{FOM} = 1 - 10 \times \left[\int |T_1(\lambda) - T_1^\star(\lambda)|^2 \mathrm{d}\lambda + \int |T_2(\lambda) - T_2^\star(\lambda)|^2 \mathrm{d}\lambda + \alpha' \int |R(\lambda)|^2 \mathrm{d}\lambda \right], \quad (2)$$

where $T_1(\lambda)$, $T_2(\lambda)$, $R(\lambda)$, and $[\cdot]^\star$ denote transmissions of output ports 1 and 2, reflection at input port at certain wavelength, and corresponding target values, respectively. To make the direct comparison between the two models, we use the same training data present in our previous work [1]. The training data are 16,000 binary HVs, which contain optimized patterns for different splitting ratios with a conventional direct binary search (DBS) or randomly generated patterns. We use the two models to generate 20 devices for four types of devices with different splitting ratios. As shown in Fig. 3, in terms of FOM, it shows a decent average improvement 2% and 5% increase of the worst case. The A-CVAE with cycle consistency clearly improves the overall generation performance.

3. Conclusion

We demonstrated a novel A-CVAE model with cycle consistency to design nano-structured splitters for an arbitrary splitting ratio. For 220 nm SOI, the design patterns generated by the new DNN model have the average total transmission (as shown in Fig. 4) of 92% across a broad bandwidth (1250 nm to 1800 nm), which is a significant improvement over our previous model without cycle consistency. Our model is generic and will be applicable to a wide range of photonic devices and metasurfaces.

References

1. Y. Tang, *et al.* “Generative deep learning model for inverse design of integrated nanophotonic devices,” Laser & Photonics Reviews (2020).
2. J. Y. Zhu *et al.*, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” ICCV, p. 2223 (2017).
3. J. Song *et al.*, “Learning to Sketch with Shortcut Cycle Consistency,” CVPR, p. 801 (2018).
4. T. Hori *et al.*, “Cycle-consistency training for end-to-end speech recognition,” arXiv:1811.01690v2 (2019).