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

Nanostructured Photonic Power Splitter Design via Convolutional Neural Networks

TaherSima, Mohammad; Kojima, Keisuke; Koike-Akino, Toshiaki; Jha, Devesh K.; Wang, Bingnan; Lin, Chungwei; Parsons, Kieran

TR2019-027 June 11, 2019

Abstract

We train a convolutional neural network (CNN) that can predict the optical response of randomly generated nanopatterned photonic power splitters in a 2 to the 400th power design space with a prediction correlation coefficient of 85%.

Conference on Lasers and Electro-Optics (CLEO)

Nanostructured Photonic Power Splitter Design via Convolutional Neural Networks

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 train a convolutional neural network (CNN) that can predict the optical response of randomly generated nanopatterned photonic power splitters in a 2^{400} design space with a prediction correlation coefficient of 85 %.

OCIS codes: 250.5300, 160.3918, 100.4996.

1. Introduction

Subwavelength scatterers can be carefully postioned on the propagation path of an incident optical beam to control its transmision and reflection wavefronts for a wide range of applications [1,2]. However, it is computationally costly to optimize a nanostructure starting from scratch every time. For example, to design photonic power divider with an arbitrary splitting ratio, photonic designers often begin with intuition or analytical models, and fine tune the structure using parameter sweep in numerical simulations after numerous iterations [3,4]. Inspired by recent progress at use of machine learning in nanophotonics [5–9], we previously used machine learning to accelerate the photonic design optimization processes using neural networks [10,11]. More recently, we developed a deep residual network to represent the design space of nanopatterned integrated photonic power splitters for inverse design [12]. Our neural network assisted nanophotonic design approaches make essential use of machine learning methods for powerful correlations between device topologies and their optical responses. To improve their performance for generalization, here we take advantage of a more powerful model called CNN (see Fig.1). Although CNN based models have been previously trained and used in nanophotonic design, a comparative study to show their generalization capability was missing.

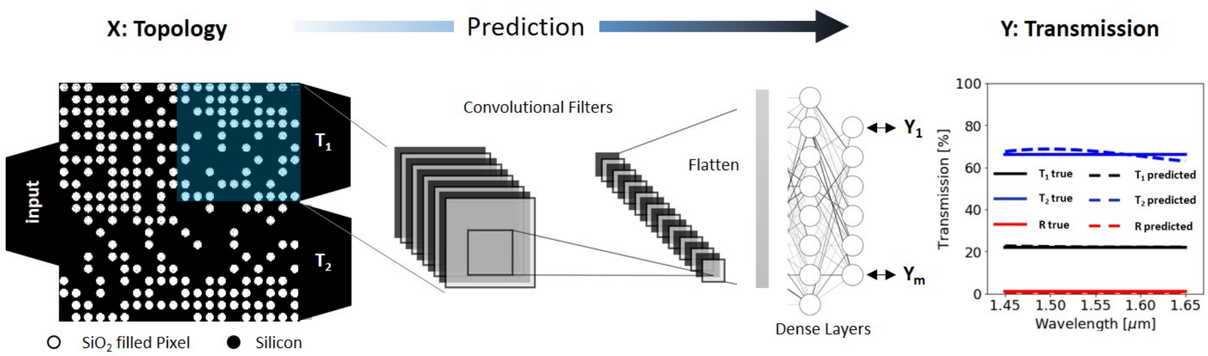

Fig. 1: Pattern recognition in nanostructured photonic devices using CNN.

2. Pattern Recognition Using Convolutional Neural networks

The goal of nanostructured power splitters is to organize optical interaction events in a compact footprint, such that the collective effect of the ensemble of scattering events guides the optical beam to a port within a target transmittance. We chose a simple three port structure on a standard fully etched silicon on insulator platform with 220 nm silicon waveguides and 2 μm buried oxide. One input and two output 0.5 μm waveguides are connected using an adiabatic taper to the 2.6×2.6 μm^2 wide square power splitter design region with a connection width of 1.3 μm. By training the model with a diverse set of semi-optimized patterns, the neural network can inverse design power splitters that outperform devices within the training data set [12]. However, we find that the implemented Fully Connected Deep Neural Networks (FCDNN) suffer from poor generalization for a large problem such as this with 2^{400} possible combinations. Meaning it will be hard for them to predict patterns that are very different from training data. Deep CNNs use convolutional filters, and thus extract better spatial features from the patterned power splitters. With this intuition, we use deep CNN network with dropout [13,14] for training our prediction model.

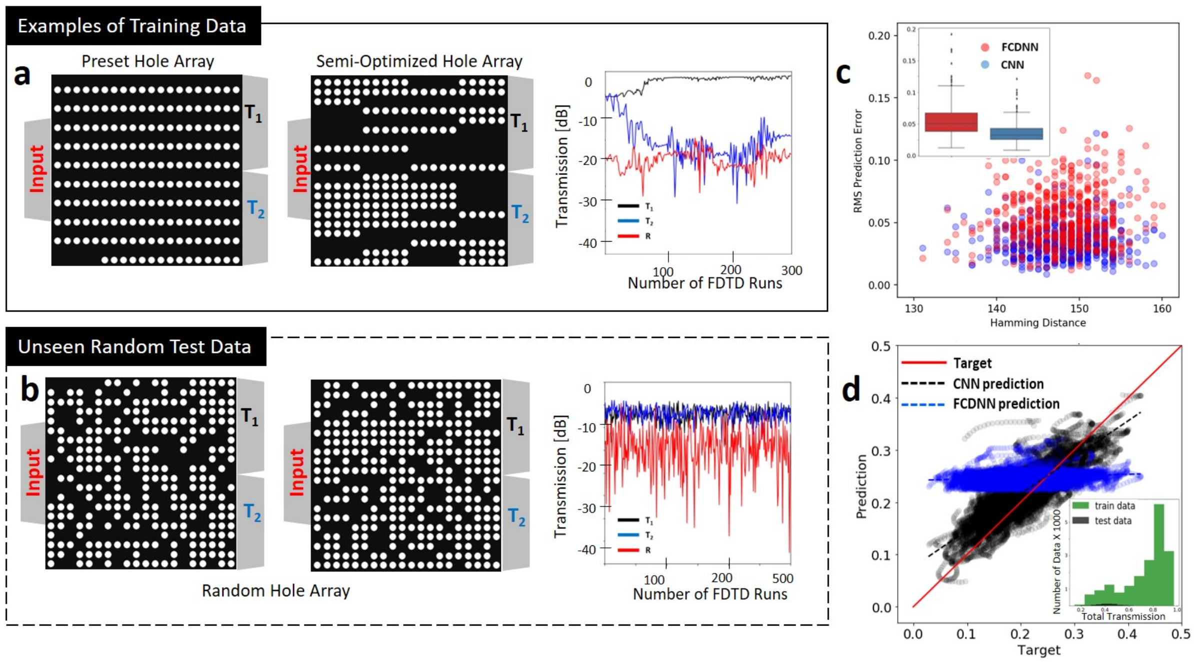

Fig. 2: a) Example of training data series and, b) randomly generated test data, c) the root mean squared (RMS) Error of prediction as a function of hamming distance of test patterns to closest train data patterns, d) correlation coefficient between target and predicted transmission spectra is 0.85 for the CNN with 3 convolutional layers and 2 dense layers and 0.16 for a 4 layers FCDNN. Inset: train and test data distributions

The architecture of our best performing network involves 3 convolutional layers with ReLU activation and 2 fully connected networks with Sigmoid activation. Using ReLU activation helps preventing vanishing gradient problem. We also find that using dropout rate of 0.4 helps to reduce overfitting. Also, local batch normalization before each convolution layer helps generalization [15].

3. Discussion and Conclusion

Deep Neural Network (DNN) models can be trained on optical response of geometrical and physical parameters of nanostructured devices. We reutilize the trained model to approximate the response of a nanostructures instantaneously. Inversely, a DNN model can suggest an approximate nanostructure for a desired response.

Additionally, we demonstrate generalization capability of a CNN based model for forward prediction of optical response. Our network achieved correlation coefficient R of up to 0.85 across the full range of transmission ratios for a randomly generated test data set.

References

1. Pendry, John B., David Schurig, and David R. Smith. "Controlling electromagnetic fields." science 312.5781 (2006): 1780-1782.
2. Faraji-Dana, Mohammad, et al. "Compact folded metasurface spectrometer." Nature communications 9.1 (2018).
3. Tian, Ye, et al. "Broadband 1× 3 Couplers With Variable Splitting Ratio Using Cascaded Step-Size MMI." IEEE Photonics J 10.3 (2018).
4. Xu, Ke, et al. "Integrated photonic power divider with arbitrary power ratios." Optics letters 42.4 (2017).
5. Liu, Dianjing, et al. "Training deep neural networks for the inverse design of nanophotonic structures." ACS Photonics 5.4 (2018).
6. Peurifoy, John, et al. "Nanophotonic particle simulation and inverse design using artificial neural networks." Science Advances 4.6 (2018).
7. Chen, Claire Lifan, et al. "Deep learning in label-free cell classification." Scientific Reports 6 (2016).
8. Asano, Takashi, et al. "Optimization of photonic crystal nanocavities based on deep learning." Optics Express 26.25 (2018).
9. Hughes, Tyler W., et al. "Training of photonic neural networks through in situ backpropagation and gradient measurement." Optica (2018).
10. Kojima, Keisuke, et al. "Acceleration of FDTD-based Inverse Design Using a Neural Network Approach.", IPR, OSA, 2017.
11. Teng, Min, et al. "Broadband SOI mode order converter based on topology optimization." OFC, OSA, 2018.
12. Tahersima, Mohammad H., et al. "Deep Neural Network Inverse Design of Integrated Nanophotonic Devices." arXiv:1809.03555 (2018).
13. Krizhevsky, Alex, et al. "Imagenet classification with deep convolutional neural networks." ADV NEUR IN. 2012.
14. Srivastava, Nitish, et al. "Dropout: a simple way to prevent neural networks from overfitting." JMLR 15.1 (2014).
15. Hinton, Geoffrey E., et al. "Improving neural networks by preventing co-adaptation of feature detectors." arXiv:1207.0580 (2012).