Diffraction-based on-chip optical neural network with high computational density

◦ Wencan Liu[1], Yuyao Huang[1], Run Sun[1], Tingzhao Fu[2] and Hongwei Chen[1]

Tsinghua University, China [1], National University of Defense Technology, China [2]

E-mail: chenhw@tsinghua.edu.cn

1. Introduction

The rapid advancement of artificial intelligence has led to substantial progress in various fields with deep neural networks (DNNs). However, complex tasks often require increasing power consumption and greater resources of electronics. On-chip optical neural networks (ONNs) are increasingly recognized for their power efficiency, wide bandwidth, and capability for light-speed parallel processing. In our previous work [1], we proposed on-chip diffractive optical neural networks (DONNs) to offer the potential to map a larger number of neurons and connections onto optics. To further improve the computational density and integration level, we proposed ultra-compact DONNs designed with the structure re-parameterization algorithm [2] and experimentally verified their performance [3], which increased the computational density by more than an order of magnitude.

2. Structure design and numerical modeling

The DONN chips developed consist of metalines, which represent the hidden layer of a neural network. These metalines consist of slots filled with SiO_2, serving as trainable parameters, optimized in the training process. To accurately model the propagation process in the metaline, slot groups and extra length between adjacent layers are utilized, resulting in the relatively large chip scale [1].

Given the challenges in directly modeling on-chip DONNs, particularly in accurately depicting the interaction between the silicon slots and the optical field, a deep complex neural network (DCNN) is employed to simulate the complex interactions in each metaline. Consequently, this approach allows for a precise numerical representation of the on-chip light propagation [2]. The architecture of the DONN with two hidden layers is illustrated in Fig. 1(a) and the modeling process is depicted in Fig. 1(b). The dimensions of the structure are set to 53 um in length and 30 um in width, with distance between adjacent layers set as 15 um. The input and output layers consist of four input straight waveguides and three output taper waveguides, respectively. As a proof of the design method, the chip was fabricated based on the SOI platform. The SEM images in Fig. 1(c) illustrates the on-chip DONN with 2 hidden layers and diffractive units. The experimental result of the chip revealing an accuracy of 93.3% in Iris plants dataset, aligning with numerical predictions [3].

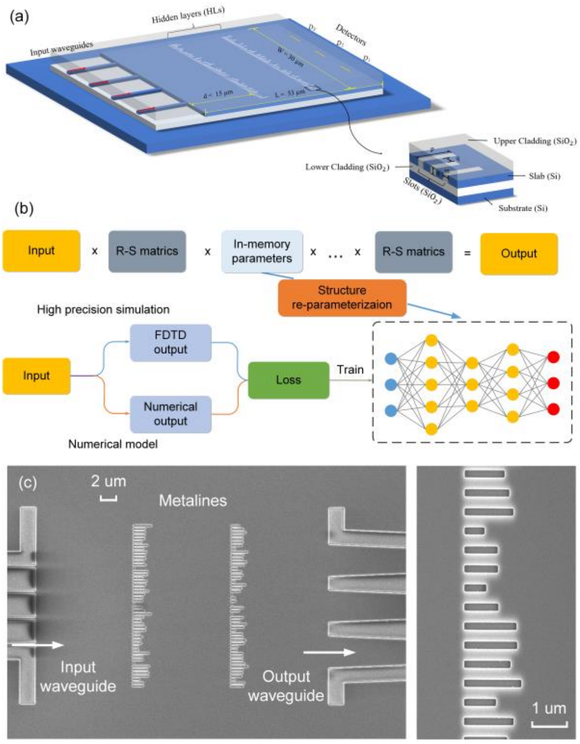

Fig. 1. (a) The architecture of the DONN. (b) Numerical modeling on structure re-parameterization algorithm. (c) SEM images of the DONN chip with 2 hidden layers and diffractive units.

3. Conclusions

In summary, DONN has significant advantage in mapping a large number of connections and trainable parameters on chip, allowing for passive computing on a compact structure. Furthermore, with utilization of the structure re-parameterization algorithm, we can greatly improve the parameter integration while achieving computational density of 18.6 $POPS/mm^2$.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (62135009).

References

[1] T. Fu et al. "Photonic machine learning with on-chip diffractive optics," Nat. Commun. 14, 70 (2023).

[2] W. Liu et al. "C-DONN: compact diffractive optical neural network with deep learning regression," Opt. Express 31, 22127–22143 (2023).

[3] W. Liu et al. "Ultra-compact optical neural network chip," CLEO 2024, no. SM4M. 5