Exploring Optical Artificial Neural Inference

Matin Dehghani[1]

[1]Abbe School of Photonics, Friedrich Schiller Universität Jena, Germany

SUMMER SCHOOL 2023

ai.uni-jena.de

INTRODUCTION

Artificial Neural Networks (ANNs) have shown immense potential in various applications, but their reliance on increasing computing power poses challenges in terms of speed and energy efficiency. To address this, researchers have been exploring alternative computing methods, and one intriguing approach is optical neural computing. **Optical neural computing** leverages the advantages of analog computing, such as **minimal energy consumption** and **intrinsic parallelism**, which can significantly accelerate computational speed. In this approach, optical signals are used to process information, and the reflection of light plays a crucial role in providing the necessary feedback mechanism for computation. **[1]**

In this poster we review methods and results from two papers including ***Diffractive Deep Neural Networks (D2NN)*** **[2]**, an all-optical deep learning framework utilizing diffractive surfaces for parallel computation, and the ***Nanophotonic Neural Medium (NNM)*** **[3]**, which employs scattering inclusions to process information.

METHODS

In a D2NN neurons are represented by individual points within the layers. The transmission coefficient of each point serves as a multiplicative bias term (Fig. 1A). Once the network is trained, local phase and height values for each point are established (Fig. 1B). These values are then utilized in generating a surface function, which enables the 3D printing of each layer (Fig. 1C). Upon printing completion, the D2NN is capable of operating at the speed of light. **[2]**

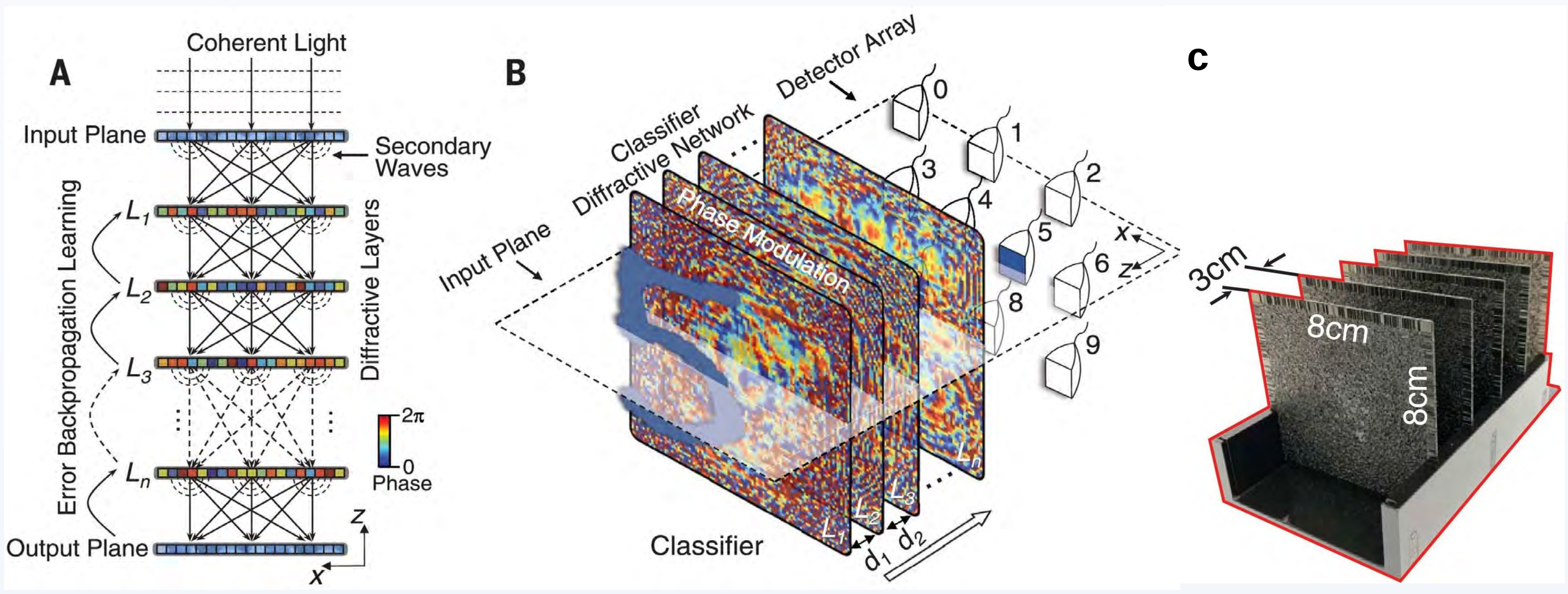

Figure 1. (A) multiple transmissive layers of D2NN. (B) Trained layers of a classifier NN. (C) 3D printed layers of the classifier. **[2]**

In nanophotonic neural networks, the nanophotonic neural medium (NNM) processes light input to generate energy distribution. Computations occurs in the host materials such as SiO2 which includes some inclusion materials (air holes or other materials) with a different refractive index than host material. These inclusions mix light similarly to matrix multiplication in digital neural networks (Fig. 2b). Inclusion positions and shapes act like weight parameters, while nonlinear functions use components like dye semiconductors for distributed activation, resembling rectified linear units (ReLUs) in traditional neural networks. **[3]**

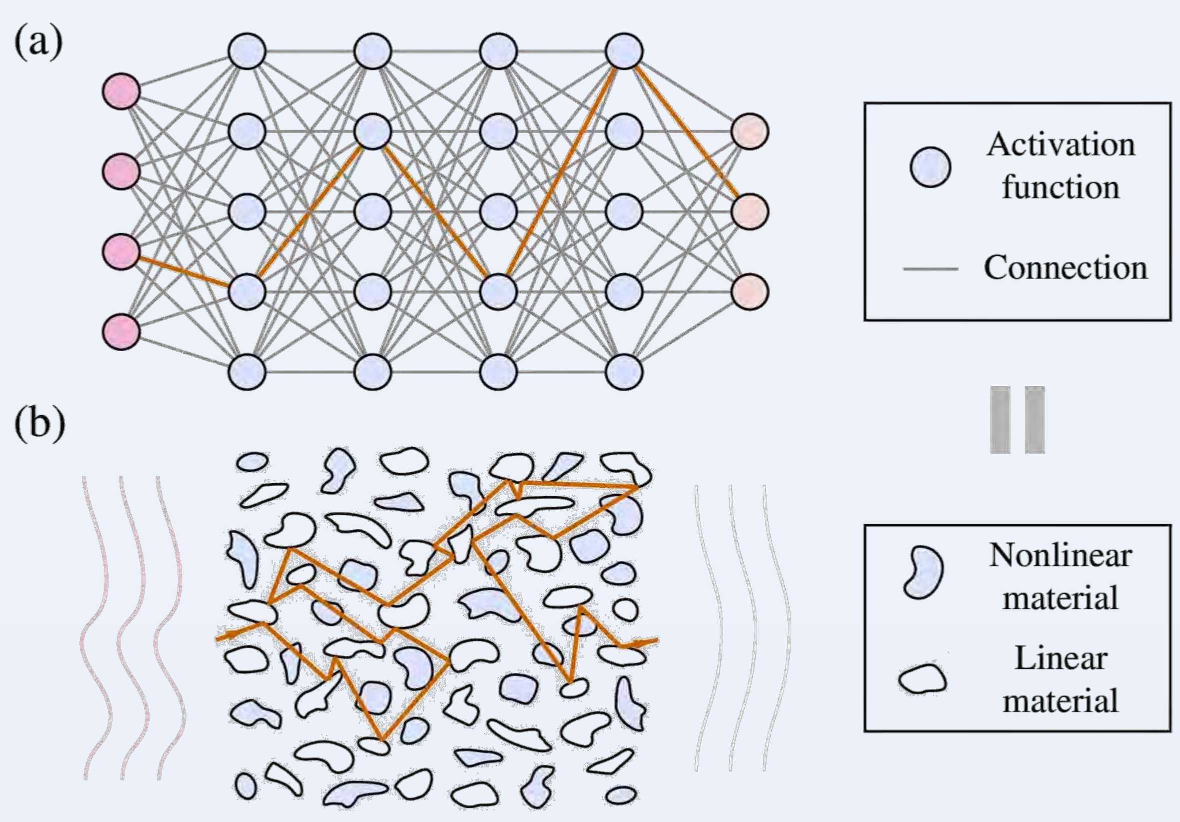

Figure 2. (a) Conventional ANN architecture. (b) Proposed Nanophotonic Neural Medium. **[3]**

RESULTS

For a five-layer D2NN, there is a 90% match between experimental and numerical results for fashion dataset with five errors out of 50 fashion products and classification accuracy of 81.13% for 10'000 samples. While it is six errors out of 50 for digit classification and 91.75% accuracy for 10'000 samples (Fig. 3&4).

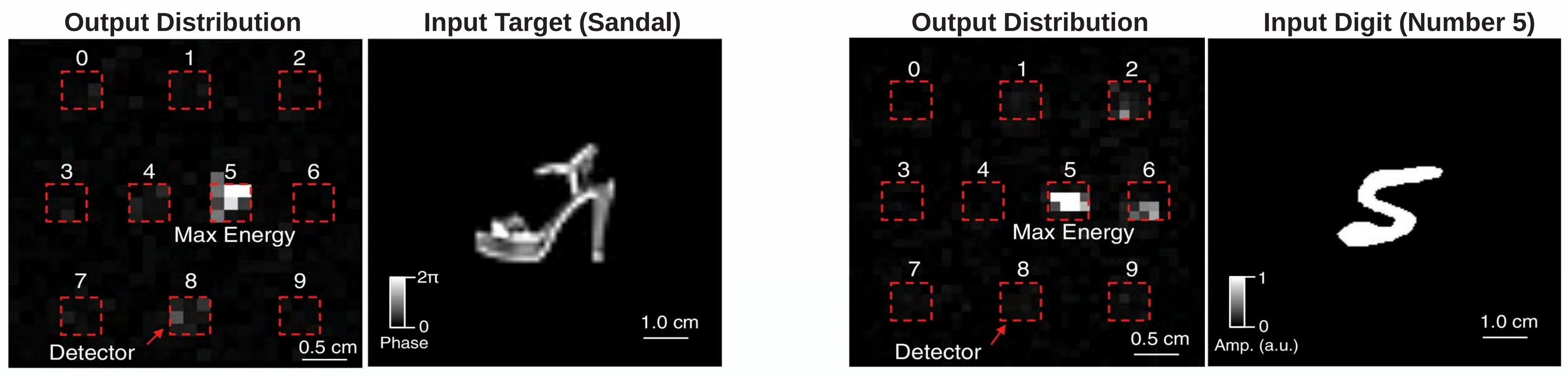

Figure 3. The input and output images of the 3D-printed D2NN for a sandal input (Fashion-MNIST class 5), and a handwritten input of "5" . **[2]**

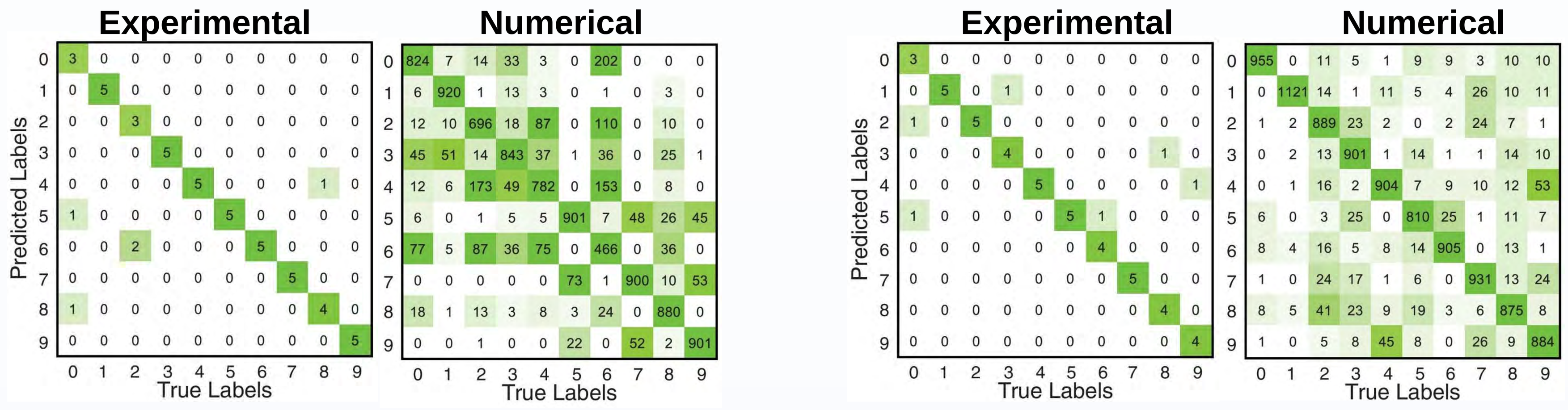

Figure 4. Confusion matrix for fashion data set (left) and digit classification (right). **[2]**

It is shown in Fig. 5 (left) that how light field is distributed inside the trained NNM for digit classification. The average recognition accuracy reaches over 79% for 1000 images of dataset. While it is possible to increase the accuracy by locating several NNMs over each other to create a stack of NNMs (Fig. 5 right). The 3D trained NNM shows accuracy of about 84% for the test set. In general, the potential number of weights can exceed 10 billion parameters per square millimeter for a 2D implementation.

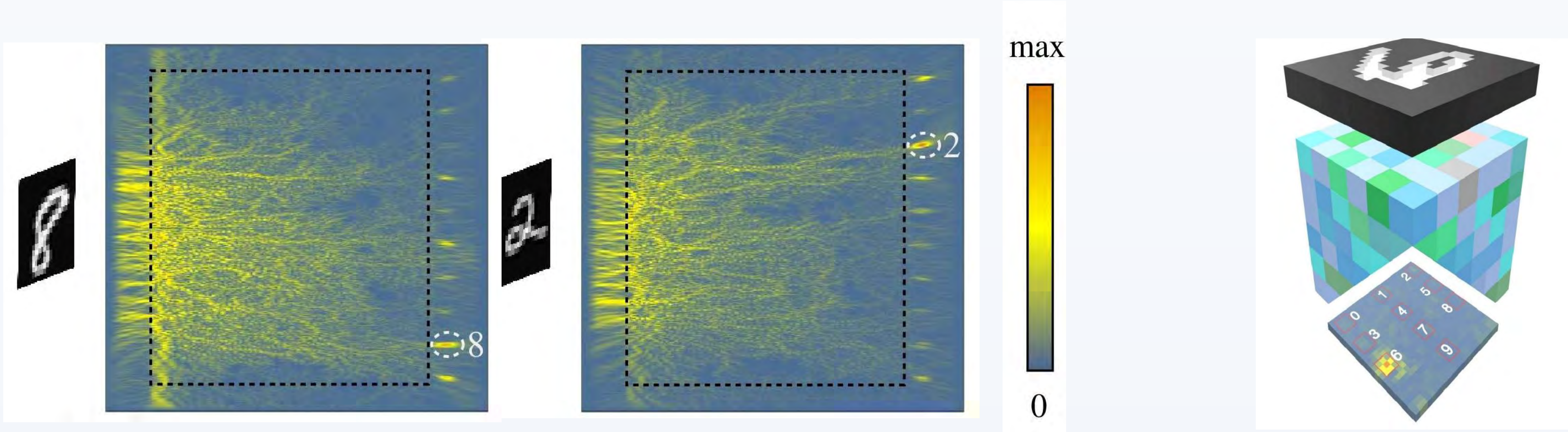

Figure 5. (left) Light field distribution in the trained NNM. (right) 3D NNM consists of multiple NNM layers. **[3]**

CONCLUSIONS

Propagation Speed of Light

Light travels at a remarkable speed of nearly 300,000 kilometers per second in a vacuum. Employing light for computation enables **ultra-fast data transmission** and processing, making it ideal for high-performance computing and real-time applications.

Energy Efficiency

Optical computing can be extremely energy-efficient compared to traditional electronic computing. Optical signals can propagate through waveguides with **minimal energy loss**, resulting in lower power consumption and reduced heat generation.

Nonlinear Light Matter Interaction

The inherent nonlinearity interaction of light and a material in Nanophotonic Neural Media serves as a **sigmoid-like activation function**, enabling distributed nonlinear activation and enhancing the capabilities of optical neural networks.

Light Wave Encoding

The wave nature of light enables us to encode information, such as images, into light waves and propagate them through a neural network. Light's ability to carry both **amplitude and phase information** simultaneously allows for precise data manipulation and complex computations within optical neural networks.

REFERENCES

[1] D. Zhang and Z. Tan, "A Review of Optical Neural Networks," Appl. Sci. 12, 5338 (2022).

[2] X. Lin et al., "All-optical machine learning using diffractive deep neural networks," Science 361, 1004-1008 (2018).

[3] E. Khoram et al., "Nanophotonic media for artificial neural inference," Photon. Res. 7, 823-827 (2019)

CONTACT

matin.dehghani@uni-jena.de