Diffractive neural networks for image processing

Viktor V. Krasnikov Egor Y. Belashov Andrey A. Grunin Andrey A. Fedyanin

Faculty of Physics of M.V.Lomonosov Moscow State University, GSP-1, Leninskie gory, Moscow 119991 Russian Federation

Abstract

Digital technologies are undergoing rapid transformation due to advancements in artificial intelligence [2], particularly deep neural networks, which are widely used in areas such as image recognition and big data processing. However, this growth increases the demand for computational resources. One promising approach to optimizing machine learning algorithms is the use of optical technologies, which can accelerate data processing due to the properties of light. This work explores the potential of optical image preprocessing in the context of spatially incoherent light, focusing on developing a model for light propagation through heterogeneous structures and applying machine learning algorithms to optimize parameters. A hybrid optoelectronic neural network was built and trained for image classification based on this model.

Motivation

Delegating computational tasks to specialized units, such as GPUs and optical processors, enables efficient processing of specific tasks with high performance and low energy consumption. Optical devices, leveraging the properties of light, provide high parallelism and low latency, making them promising for solving tasks that require intensive computations.

Research objectives

Theory

A diffractive neural network represents an array of optically heterogeneous thin masks. On each of these masks, the phase and/or amplitude of the field is modulated point by point. The system is divided into an initial plane, where the image is projected, the planes after the masks, and the output plane. This system is associated with a perceptron-type neural network because the field at each point of a plane is linearly related to the fields at points in the previous plane. An important distinction is that the number of transition weights between layers is much smaller [3]. This is due to the fact that the field at any point in the plane after the mask is calculated as an integral of the multiplication of the field in the previous plane with a fixed kernel, followed by multiplying this value by a controlled coefficient. The mathematical description is presented in Equation below:

$$u_i(x,y) = w_i(x,y)\iint_{-\infty}^{+\infty} u_{i-1}(\xi,\eta)\frac{e^{ik\sqrt{L^2+(x-\xi)^2+(y-\eta)^2}}}{\sqrt{L^2+(x-\xi)^2+(y-\eta)^2}}\,d\xi\,d\eta \tag{1}$$

In this equation, u_i, u_{i-1} represent the field in the current and previous layers, respectively, L is the distance between the layers, and $w_i(x,y)$ is the weight of the neuron at the point (x,y). The number of neurons and the discretization of this equation depend on the degree of accuracy with which optical inhomogeneities can be created. Thus, for a transition between layers with N neurons, a perceptron has N^2 weights, whereas a diffractive neural network has N.

Method of random phase generation of the initial field: A random Fourier pattern is generated, where the value at each point is equally likely to range from 0 to 1, and is constrained by a Gaussian function with a given dispersion $\frac{1}{\sigma^2}$.

Calculation of coherent and incoherent propagation: Free-space wave propagation methods aim at solving the homogeneous Helmholtz equation [1]:

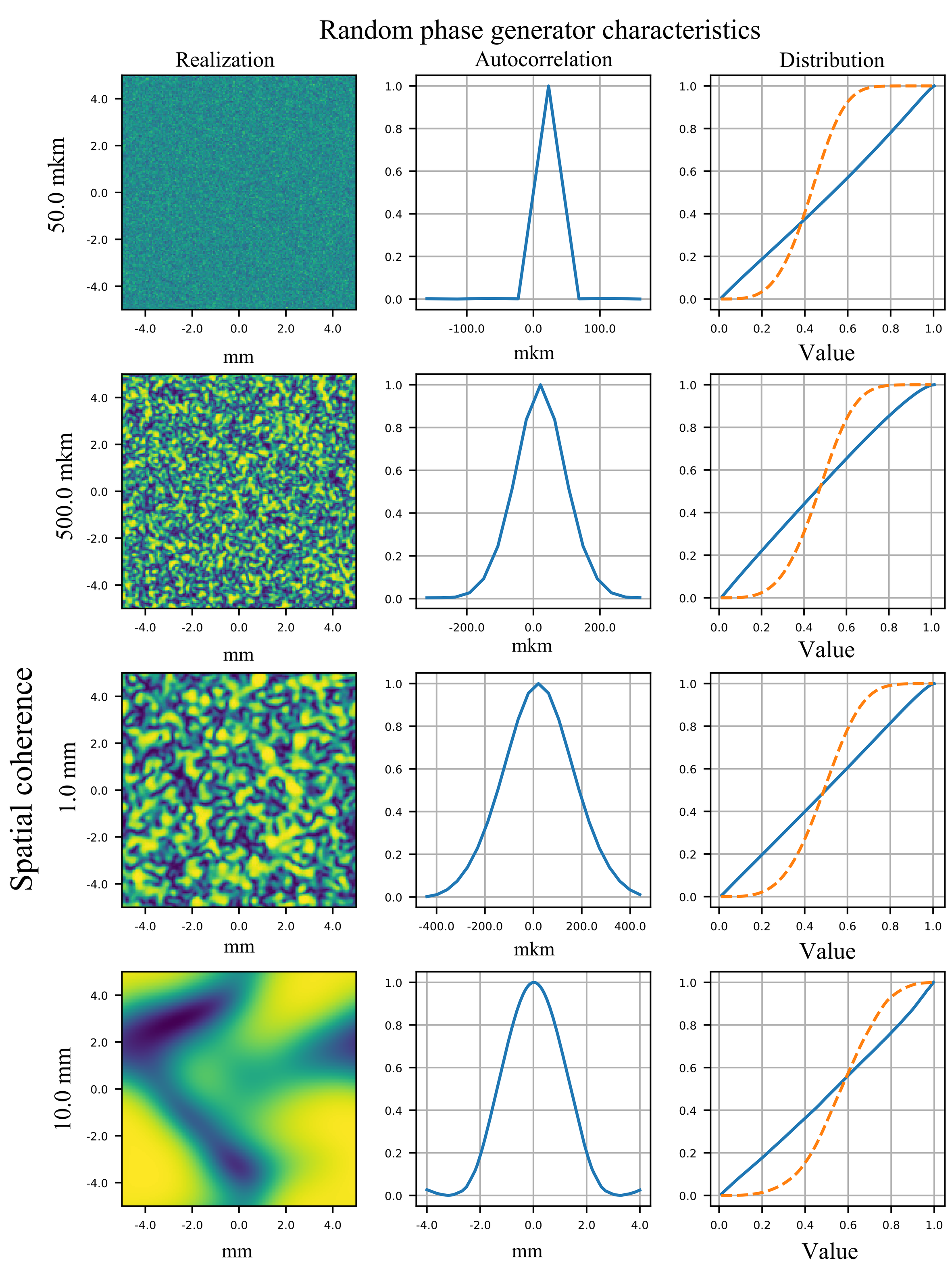
$$\nabla\psi + k^2\psi = 0 \tag{2}$$

Mathematically, AS propagation can be formulated as:

$$\psi = \mathcal{F}^{-1}[H_{AS}\mathcal{F}[\psi_0]] \tag{3}$$

Transfer function:

$$H_{AS}(f_x, f_y) = e^{i2\pi z\sqrt{\frac{1}{\lambda^2}-f_x^2-f_y^2}} \tag{4}$$



Results and discussion

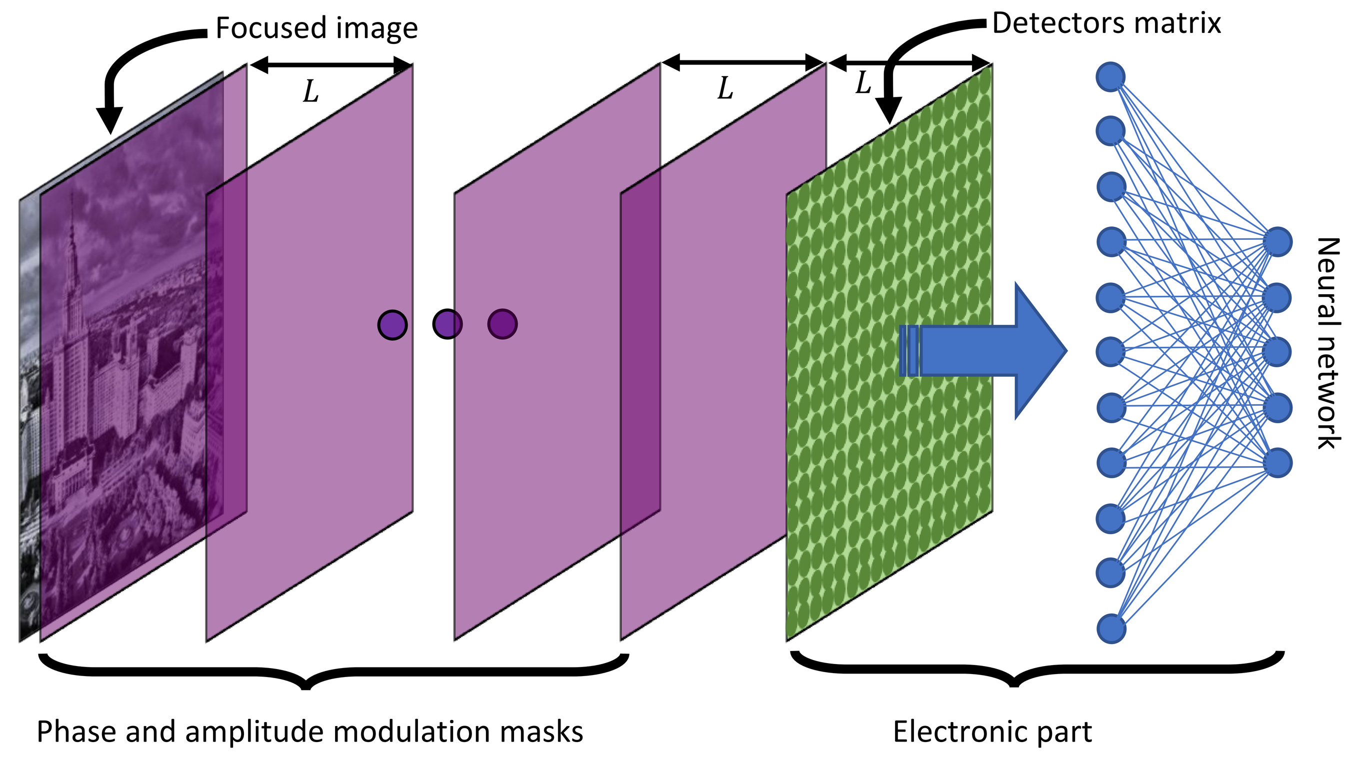


Figure 1. Schematic diagram of optoelectronic neural network.

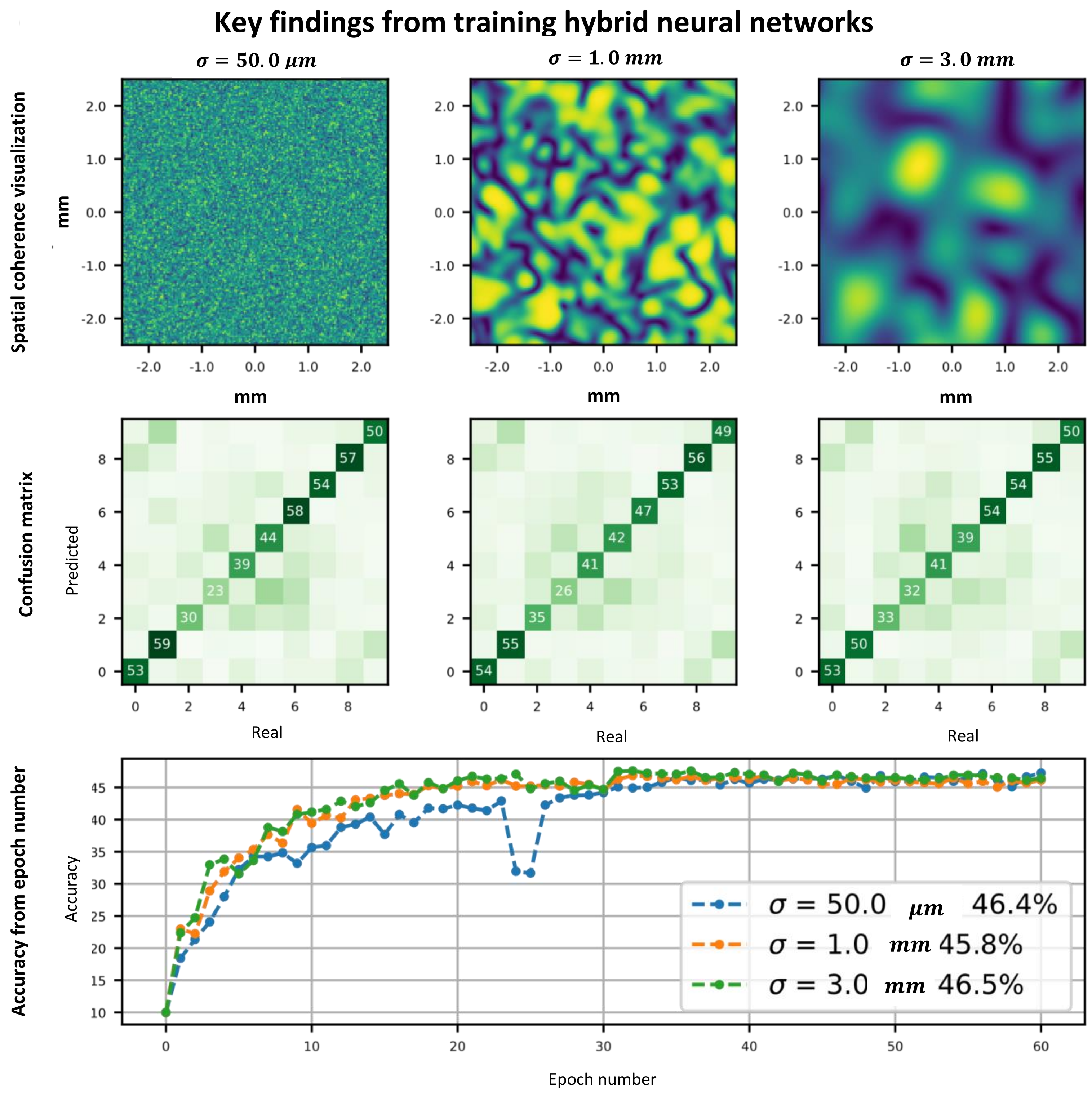


Figure 2. Training results of hybrid neural networks. The first row - visualisations of a particular realisation of the incoherent field phase. Second row - error matrix: colour represents the percentage of answers, numbers on the diagonal represent the percentage of correct answers. The third row is a graph of the dependence of accuracy on the epoch number for three models.

Conclusions

By training optoelectronic networks for three levels of spatial coherence ($50\mu m, 1000\mu m, 3000\mu m$), it is shown that the models give the same results for a large number of epochs, regardless of the degree of coherence. The small accuracy for some classes is explained by the similarity of the images. Phase and amplitude modulation patterns confirm that optical layer training did occur. Examples of the networks show that the system performs image transformations rather than simply focusing the image on the detectors.

References

[1] Rainer Heintzmann, Lars Loetgering, and Felix Wechsler. Scalable angular spectrum propagation. *Optica*, 10(11):1407–1416, 2023.

[2] Jiamin Wu, Xing Lin, Yuchen Guo, Junwei Liu, Lu Fang, Shuming Jiao, and Qionghai Dai. Analog optical computing for artificial intelligence. *Engineering*, 10:133–145, 2022.

[3] Tao Yan, Jiamin Wu, Tiankuang Zhou, Hao Xie, Feng Xu, Jingtao Fan, Lu Fang, Xing Lin, and Qionghai Dai. Fourier-space diffractive deep neural network. *Physical review letters*, 123(2):023901, 2019.

Website

krasnikovvv@my.msu.ru