OPTIMIZATION OF NANOSTRUCTURED FREE-FORM OPTICAL FIBERS SUPPORTING THREE WEAKLY-COUPLED SPATIAL MODES USING DENSE AND CONVOLUTIONAL NEURAL NETWORKS

[1]Bartosz PALUBA, [2]Maciej NAPIORKOWSKI, [1,3]Rafal KASZTELANIC, [1]Ryszard BUCZYNSKI

[1]University of Warsaw, Faculty of Physics, Pasteura 5, 02-093, Warsaw, Poland, EU, bartosz.paluba@fuw.edu.pl, rafal.kasztelanic@fuw.edu.pl, ryszard.buczynski@fuw.edu.pl.

[2]Wroclaw University of Science and Technology, Faculty of Fundamental Problems of Technology, Wybrzeze Wyspianskiego 27, 50-370, Wroclaw, Poland, EU, maciej.napiorkowski@pwr.edu.pl.

[3]Lukasiewicz Research Network, Institue of Microelectronics and Photonics, Wolczynska 133, 01-919, Warsaw, Poland, EU, Rafal.Kasztelanic@imif.lukasiewicz.gov.pl.

https://doi.org/10.37904/nanocon.2025.5167

Abstract

The ever-growing need for increasing bandwidth of telecommunication networks forces engineers and scientists to develop new solutions, concepts and devices for data transfer. This project aims to determine the best possible structures of few-mode optical fibers for Mode-Division Multiplexing using neural networks of different kinds. Applying the Generative Inverse Design Networks (GIDNs) approach, we studied optimization of the three-mode fibers composed of two types of glass rods: pure SiO_2 and SiO_2 doped with 5% of GeO_2. We compared principles of operation and efficiency of the dense neural network (DNN) and the convolutional neural network (CNN) within two cycles of GIDNs. We observed that convolutional network outperformed the dense one just after first cycle achieving higher average as well as maximal values of minimal separations between modes. The best structures obtained with the DNN were characterized by minimal separations exceeding 1.98×10^{-3}, while the best fiber optimized with the CNN had a minimal separation of ca. 2.11×10^{-3}. Moreover, the latter was applied for the first time for the free-form fibers optimization and obtained results encourage to continue research using the CNN as a main numerical tool.

Keywords: Few-Mode Fibers, Convolutional neural network, Dense neural network, Free-Form Nanostructured optical fibers

1. INTRODUCTION

Simultaneous transferring of multiple signals in a single medium, such as an optical fiber, is possible due to multiplexing techniques [1]. These signals (referred to as channel) can be combined into one signal, transmitted through shared medium and divided back into original form using different properties of electromagnetic radiation. One of such techniques is the Space-Division Multiplexing (SDM) [2]. It can be realized in several ways: a) bundles of single-mode fibers with reduced cladding thickness, b) multicore fibers (MCFs) with independent, single-mode cores, c) coupled-cores fibers or d) few-mode fibers (FMFs) of a single core [3]. In our research we are focused on the last group of fibers in which supported spatial modes are weakly-coupled, i.e. a difference between effective refractive indices of adjacent modes ($|\Delta n_{eff}|$) is bigger than 1×10^{-3} [4]. It allows transferring multiple signals through single FMF in which each spatial mode acts as an independent channel or as two coupled channels, if polarization degeneracy is counted. Unlike in the case of strongly-coupled FMFs [4-6], weak coupling between adjacent modes enables to apply much cheaper telecommunication systems and reduce costs due to less complicated digital analyzes of receiving signals [7].

It was shown that separation between groups of modes is enhanced, if the core of an optical fiber contains several concentric rings of different refractive indices [8-11]. Since the shapes, sizes and positions of extrema

of electric field are different for certain groups of modes, the influence of the medium of particular ring would be stronger on one mode and much weaker on another. However, spatially degenerated modes (e.g. LP_{11a}, LP_{11b}) would interact with such rings exactly in the same way. Hence, to separate them, one should introduce additional modulation of refractive index within each ring. Fabrication of fibers with such complex internal structure is impossible with traditional methods, e.g. using vapor deposition methods [12, 13]. Instead, one can use the nanostructurization technology in which thousands of glass rods of two or more types are stacked in a designed pattern. Due to this approach one can create arbitrary distribution of refractive index in a core to achieve desired properties of an optical fiber referred to as free-form nanostructured fiber.

Having 4000 rods of two types, one would need to consider 2^{4000} possible structures of a fiber's core. Such complexity of the system requires advanced numerical tools involving statistical analysis to find optimal structure for certain application. Promising approach is to use various kinds of neural networks. In previous research we have shown that using the dense neural networks within the Generative Inverse Design Networks (GIDNs) framework one can efficiently optimize the structure of ten-mode free-form nanostructured fiber with minimal separation $\mathrm{Min}|\Delta n_{eff}|$ between modes equal to 1.2×10^{-3} [14]. In this work we have focused on the simpler case of fibers supporting three spatial modes. We have simultaneously applied the dense (DNN) and the convolutional (CNN) neural networks and we have compared their performance during two GIDNs cycles. It is worth mentioning that the latter network was tested for the first time in the context of optimization of free-form nanostructured optical fibers.

2. METHODS

In this research, the Generative Inverse Design Networks (GIDNs) approach [15] was applied to design the structure of the free-form nanostructured fiber's core which can guide three spatial modes (corresponding to LP_{01}, LP_{11a} and LP_{11b}) with as high separations $|\Delta n_{eff}|$ as possible. The idea of this approach is based on cycles of neural networks training and successive optimization process. Obtained dataset is verified and examples fulfilling certain criteria are appended to the initial learning dataset. Afterwards the whole cycle is repeated until the desired outcome is achieved. Randomly generated fiber are mainly of poor and moderate properties, hence this methodology allows enriching the training datasets with better, near-optimal examples. Consequently, with each cycle the network's accuracy towards structures of the best performance would increase.

Network's trainings and optimizations were performed using Keras 3.0 from TensorFlow platform [16]. Predictions of neural networks were verified by conducting calculations using BPM-Matlab – an open-source tool for optical calculations within Matlab software [17].

2.1 Initial learning dataset

The investigated fiber cores consisted of five concentric rings, where the smallest one was a circle treated as a ring of inner radius equal to zero. Four rings had a randomly chosen outer diameter, while the fifth, outermost one had a constant outer diameter equal to 13.86 μm. Moreover, these rings had a periodically modulated refractive index distribution described by the formula:

$$n_{loc,i}(\theta) = n_{av,i} \cdot [1 + \delta_i \cos(M_i \theta)] \quad (1)$$

where:

$n_{loc,i}$ – local refractive index of i-th ring,

θ – angle with respect to a center of a core (rad),

$n_{av,i}$ – average refractive index of an i-th ring,

$|\delta_i|$ – amplitude of modulation

M_i – number of maxima/minima of refractive index on i-th ring

Values of n_{av} and δ parameters were continuous and adjusted in such a way that the lowest possible value of n_{loc} was 1.44442 which corresponded to pure SiO_2 glass at 1550 nm and the highest one was equal to 1.45193 corresponding to SiO_2 glass doped with 5% mol of GeO_2 at 1550 nm as well. To sum up, each five-ring core was described by a set of 19 parameters (4 outer diameters, 5 n_{av} values, 5 amplitudes δ and 5 numbers M) of values randomly chosen from a specified sets and one constant value of a total diameter.

Even such complicated fiber's cores with rings of modulated refractive indices can be fabricated using the idea of nanostructurization, where the continuous distribution given by equation (1) is replaced by a discrete structure composed of thousands of rods. Hence, each fiber's core described by the 19 parameters was created based on the lattice consisted of 40 layers of hexagonal glass rods, which corresponds to 4921 rods in total. There were two types of rods: made of pure SiO_2 glass and of SiO_2 glass doped with 5% mol of GeO_2, corresponding to assumed maximal and minimal refractive index values. The distance between the nearest rods (lattice constant) was equal to 0.2 μm. To obtain the final, discrete structure of the core, the circle was inscribed into the 40-layer lattice and excessive, outer rods were removed, which explains the value of total diameter of 13.86 μm.

Then, effective refractive indices (n_{eff}) of modes supported by each core at 1550 nm were calculated using BPM-MATLAB software based on figures of refractive index distributions. Afterwards, fibers having third mode of a n_{eff} value above 1.445 and fourth mode below 1.445 were selected from the whole collected dataset. The threshold value was chosen based on the observation that fourth modes of n_{eff} < 1.445 exhibited increased losses, hence corresponding fibers were treated as three-mode structures and the forth mode was neglected.

2.2 Neural networks architecture and training step

The dense neural network (DNN) was composed of four identical layers of 512 neurons each with a rectified linear unit (ReLU) as an activation function, whereas the fifth, output layer was activated using linear function. Input used by the DNN was composed of vectors of a size 1×49, which were obtained from 19 values describing each fiber's core. Rings' radii, average refractive indices n_{av} and amplitudes δ were normalized to the values from the [0, 1] set, while M numbers were transformed into sequences of zeros and ones using one hot encoding. The output consisted of vectors of n_{eff} values of three modes normalized to the values from the [0, 1] set.

On the other hand, the convolutional neural network (CNN) consisted of three 2D convolutional layers of 32, 64 and 128 neurons of 3×3 kernels activated by the ReLU function, which were separated by 2D average pooling layers of 2×2 kernels. Next, two dense layers of 256 neurons and the output layer, activated by the ReLU and the linear function respectively, were added. In this case, the input consisted of figures of fibers cross-sections converted into 87×87 square matrices of zeros (pure silica glass rods) and ones (doped silica glass rods). As in the case of the DNN, the output consisted of vectors of three neff values normalized to the [0, 1] set.

Both neural networks were trained in an analogous manner in each GIDNs cycle. Prepared datasets were split randomly at a ratio of 80:20 into training and validation sets. To monitor the training process, in which the predicted values (outputs) were continuous, a mean square error (MSE) was used as a loss function and a mean absolute error (MAE) was an additional metric. It is worth recalling, that although the training procedures were similar, the inputs needed to be prepared in an entirely different way for the DNN and for the CNN.

2.3 Random structures optimization step

The set of structures for optimization was randomly generated in an analogous way as for initial training datasets, but the values of effective refractive indices as outputs weren't necessary, because networks relied solely on their own predictions. Since the form of data and the operation principles of DNN and CNN are very distinct, the process of optimization had some major differences as well. In the case of the DNN, 100 thousand

structures were optimized at once, while the CNN operated with 1000 to 2500 structures due to much higher RAM usage during the process. The optimization was carried out for 10 thousand steps and it was performed based on customized loss functions. At the end of each cycle, 10 thousand structures verified with BPM-MATLAB meeting the requirement on $n_{eff,3}$ and $n_{eff,4}$ as for the initial dataset were appended to the learning dataset.

3. RESULTS AND DISCUSSION

Initial learning datasets and the details concerning training and optimization were adjusted in such a way, that the comparison of applied neural networks, despite many differences, would be reliable. One can analyze the time consumption and CPU or GPU load during the operation. Since the input data of the DNN consisted of relatively short vectors the corresponding mathematical operations were computationally simple and nondemanding with respect to e.g. RAM usage. That is why during the optimization steps, 100 thousand structures (or even 200 thousand) could be processed at once. However, randomly chosen values *M* of extrema on each ring and number of rings were constant during whole optimization and the network was adjusting only the size of each ring, average refractive indices and the amplitudes of modulation. Consequently, many cores had unfavorable structure with respect to the field distribution of considered three modes. These structural drawbacks often couldn't be overcame during optimization, which resulted in many poor fibers. Therefore, only 5% of the best optimized structures were selected for verification and ca. 30% of verified ones met the specified requirements on effective refractive indices.

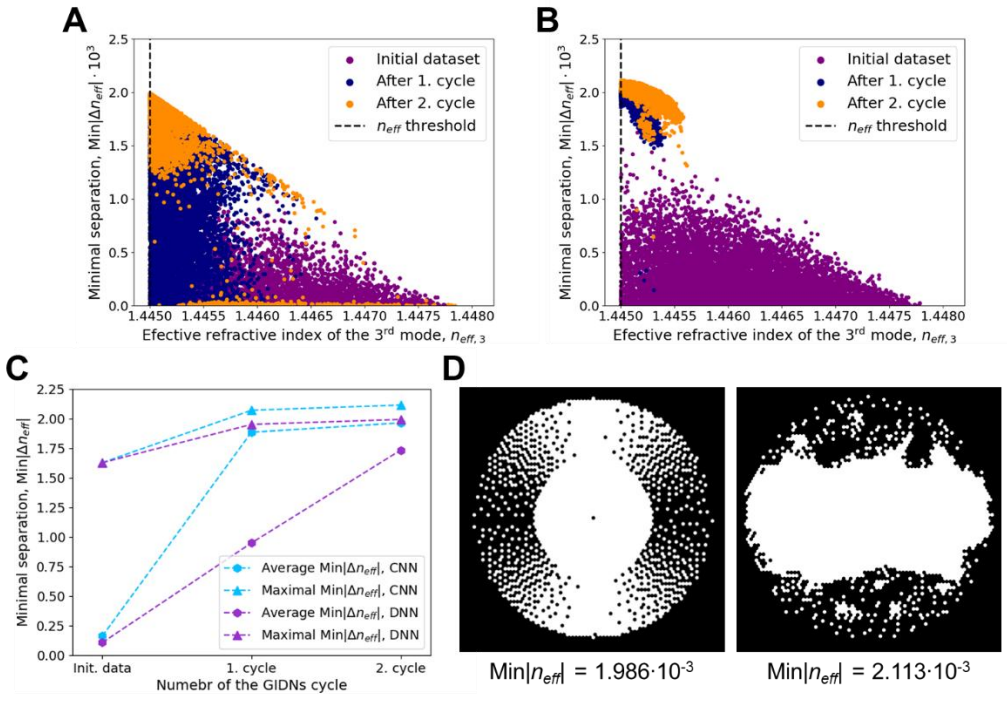

Figure 1 Distributions of structures in initial learning datasets and after the first and the second GIDNs cycle with the DNN (A) and the CNN (B). Comparison of average and maximal Min|Δn_{eff}| (C) and cross-sections of the best structures (D) obtained with the DNN (left) and the CNN (right) after the second cycle.

On the other hand, the computations with the CNN were much more demanding and required constant usage of GPU. Instead of vectors, the network operated on matrices of dimensions 87×87. Additionally, the network was composed of more layers: three convolutional, two pooling and two dense layers with respect to four dense layers in the DNN. It remarkably extended the time necessary for calculations and to avoid RAM overload, only 1000 or 2000 fibers could have been optimized simultaneously. Nevertheless, the CNN was

able to rearrange rods of the core without any limitations, which caused that all of obtained fibers were worth verifying and most of them (60 – 90%) could be appended to the learning dataset after verification.

To have better insight into the effectiveness of used neural networks, one should analyze data obtained with each network after verification step. Distributions of fiber structures in the initial datasets and structures obtained due to first and second GIDNs cycle are presented in **Figure 1A** and **B**. It is convenient to analyze relation between Min|Δn_{eff}| and n_{eff} of the last mode of interest (in this case the third mode). In general, the lowest the value of the third n_{eff} is the higher the possible separation, so the random examples within initial sets exhibit the triangular distribution. However, major changes appear just after first cycle. Based on very similar initial learning dataset, the DNN designed fibers of $n_{eff,3}$ shifted towards lower values and found many examples of higher Min|Δn_{eff}|, but the CNN was able to optimize almost all structures to have very high separations. After the second GIDNs cycle, the DNN found much better structures than after the first one, but still it was undoubtedly outperformed by the CNN.

Above observations are additionally emphasized in **Figure 1C**. Both networks started with the learning sets of comparable average Min|Δn_{eff}|, but the first sets of verified structures had average values of 0.95×10^{-3} and 1.88×10^{-3} in the case of the DNN and the CNN, respectively. After the second cycle the difference between networks significantly decreased. In the case of structures having the highest separations, the CNN also outperformed the DNN. Examples of the best fibers together with their Min|Δn_{eff}| obtained with both networks are depicted in **Figure 1D**. Clear structural differences between these two fiber are visible. As it was mentioned, the DNN was limited to rings-assisted cores, however to obtain very high Min|Δn_{eff}| of 1.986×10^{-3} only two rings out of five were enough, apparently. The middle of the core had high, homogeneous refractive index equal to the maximal value of 1.45193 and was surrounded by the ring of M = 2 with moderate modulation of the n_{av} value. The remaining three rings had negligibly small thickness and can't be seen on the figure.

On the other hand, the CNN radically modified the ring-assisted structure and no rings can be observed in the optimized structures of the highest separation. The fiber shown in **Figure 1D** achieved the Min|Δn_{eff}| equal to 2.113×10^{-3} although its structure was highly asymmetric. Such irregularities may affect optical properties of the resultant fibers, e.g. enhance bending and twisting losses or cause excessive mechanical stress around the core. In this case the additional symmetrizing of the core's structure would be beneficial.

4. CONCLUSION

In this paper, for the first time the application of convolutional neural networks for optimization of free-form nanostructured fiber was presented and compared with the dense neural network. Both networks were trained in two cycles according to the Generative Inverse Design Networks approach. We observed that the CNN outperformed the DNN significantly just after a first training – optimization process and designed much better three-mode fibers. After the second GIDNs cycle, examples of the best structures obtained with the DNN and the CNN achieved minimal separations between modes equal to 1.986×10^{-3} and 2.113×10^{-3}, respectively. Differences between networks arose from distinct principles of operation. The DNN was limited to cores with five rings and once chosen set of number of extrema on each ring remained unchanged during whole optimization. On the other hand, the CNN could have arbitrarily rearranged the rods within the core. Nevertheless, both approaches resulted in very promising results and should be farther developed upon usage for free-form nanostructured fibers design.

ACKNOWLEDGEMENTS

This work was supported by the National Science Center, Poland [grant number MAESTRO 14, UMO-2022/46/A/ST7/00238]

REFERENCES

[1] MOHAMMADY, S. *Multiplexing - Recent Advances and Novel Applications.* IntechOpen, 2022. DOI:10.5772/intechopen.95684

[2] NAKAZAWA, M., SUZUKI, M., AWAJI, Y., MORIOKA, T. *Space-Division Multiplexing in Optical Communication Systems: Extremely Advanced Optical Transmission with 3M Technologies*. Springer Cham, 2022. DOI:10.1007/978-3-030-87619-7.

[3] RICHARDSON, D.J., FINI, J.M., NELSON, L.E. Space-division multiplexing in optical fibres. *Nature Photonincs*. 2013, vol. 7, pp. 354–362. DOI:10.1038/nphoton.2013.94.

[4] SILLARD, P., MOLIN, D. A Review of Few-Mode Fibers for Space-Division Multiplexed Transmissions. In: *39th European Conference and Exhibition on Optical Communication (ECOC 2013)*. London: Institution of Engineering and Technology, 2013, pp. 1–3. DOI:10.1049/cp.2013.1272.

[5] SAKAMOTO, T., MORI, T., WADA, M., YAMAMOTO, T., YAMAMOTO, F., NAKAJIMA, K. Strongly-coupled multi-core fiber and its optical characteristics for MIMO transmission systems. *Optical Fiber Technology*, 2017, vol. 35, pp. 8–18. DOI:10.1016/j.yofte.2016.07.010.

[6] SILLARD, P., BIGOT-ASTRUC, M., MOLIN, D., Few-Mode Fibers for Mode-Division-Multiplexed Systems. *Journal of Lightwave Technology*. 2014, vol. 32, pp. 2824–2829. DOI:10.1109/JLT.2014.2312845.

[7] SU, Y., HE, Y., CHEN, H., LI, X., LI, G. Perspective on mode-division multiplexing. *Applied Physics Letters*. 2021, vol. 118, 200502. DOI:10.1063/5.0046071.

[8] JIANG, S., MA, L., ZHANG, Z., XU, X., WANG, S., DU, J., YANG, C., TONG, W., HE, Z. Design and Characterization of Ring-Assisted Few-Mode Fibers for Weakly Coupled Mode-Division Multiplexing Transmission. *Journal of Lightwave Technology*. 2018, vol. 36, pp. 5547–5555. DOI:10.1109/JLT.2018.2874526.

[9] GE, D., GAO, Y., YANG, Y., SHEN, L., LI, Z., CHEN, Z., HE, Y., LI, J. A 6-LP-mode ultralow-modal-crosstalk double-ring-core FMF for weakly-coupled MDM transmission. *Optics Communications*. 2019, vol. 451, pp. 97–103. DOI:10.1016/j.optcom.2019.06.015.

[10] HE, Z., DU, J., CHEN, X., SHEN, W., HUANG, Y., WANG, C., XU, K., HE, Z. Machine learning aided inverse design for few-mode fiber weak-coupling optimization. *Optics Express*. 2020, vol. 28, 21668. DOI:10.1364/OE.398157.

[11] WANG, Z., LU, Q., TU, J., XIAO, Q., SHEN, L., LAN, X., LI, Z., YU, C., LU, C. Design, fabrication, and characterization of a low-index center and trench-assisted 7-ring-core 5-mode-group fiber for dense space-division multiplexing. *Optics Express*. 2022, vol. 30, pp. 650 – 663. DOI:10.1364/OE.447823.

[12] WANG, X., NIE, Q., XU, T., LIU L. A review of the fabrication of optic fiber. In: *ICO20: Optical Design and Fabrication*. Proceedings of the SPIE, 2006, pp. 346-354. DOI:10.1117/12.668147.

[13] DHAR, A., CHOUDHURY, N., SEN, R. Advancements in fabrication technology for rare-earth doped optical fibers: A retrospective analysis. *Optics Communications*. 2025, vol. 577, 131463. DOI:10.1016/j.optcom.2024.131463.

[14] NAPIORKOWSKI, M., KASZTELANIC, R., BUCZYNSKI, R. Optimization of spatial mode separation in few-mode nanostructured fibers with generative inverse design networks. *Engineering Applications of Artificial Intelligence*. 2024, vol. 133, 107955. DOI:10.1016/j.engappai.2024.107955

[15] CHEN, C., GU, G.X. Generative Deep Neural Networks for Inverse Materials Design Using Backpropagation and Active Learning. *Advanced Science.* 2020, vol. 7, 1902607. DOI:10.1002/advs.201902607

[16] TensorFlow. *Keras: The high-level API for TensorFlow.* [online]. 2023. [viewed: 2025-04-23]. Available from: https://www.tensorflow.org/guide/keras?hl=pl.

[17] VEETTIKAZHY, M., HANSEN, A.K., MARTI, D., JENSEN, S.M., BORRE, A.L., ANDRESEN, E.R., DHOLAKIA, K., ANDERSEN, P.E. BPM-Matlab: an open-source optical propagation simulation tool in MATLAB. *Opt. Express*. 2021, vol. 29, 11819. DOI:10.1364/OE.420493.