Ultra-large mode area multi-core orbital angular momentum transmission fiber designed by neural network and optimization algorithms*

GU Zhiwei[1,2], HUANG Wei[1,2], ZHANG Ran[1,2], FAN Junjie[1,2], and SONG Binbin[1,2]**

1. Key Laboratory of Computer Vision and Systems (Ministry of Education), School of Computer Science and Engineering, Tianjin University of Technology, Tianjin 300384, China

2. Engineering Research Center of Learning-Based Intelligent System (Ministry of Education), Tianjin University of Technology, Tianjin 300384, China

(Received 13 March 2023; Revised 12 May 2023)

A large mode area multi-core orbital angular momentum (OAM) transmission fiber is designed and optimized by neural network and optimization algorithms. The neural network model has been established first to predict the optical properties of multi-core OAM transmission fibers with high accuracy and speed, including mode area, nonlinear coefficient, purity, dispersion, and effective index difference. Then the trained neural network model is combined with different particle swarm optimization (PSO) algorithms for automatic iterative optimization of multi-core structures respectively. Due to the structural advantages of multi-core fiber and the automatic optimization process, we designed a number of multi-core structures with high OAM mode purity (>95%) and ultra-large mode area (>3 000 μm^2), which is larger by more than an order of magnitude compared to the conventional ring-core OAM transmission fibers.

Document code: A **Article ID:** 1673-1905(2023)12-0744-8

DOI https://doi.org/10.1007/s11801-023-3046-5

Orbital angular momentum (OAM) modes, which are promising to improve the capacity of fiber communication systems, have attracted extensive research interests in recent years[1-3]. Due to the unique optical properties of OAM modes, conventional fibers cannot support stable transmission of OAM states and the near degeneracy of vector eigenmodes is easy to occur. Accordingly, various kinds of OAM transmission fibers with high contrast-index ring-core structures are emerging in an endless stream[4-7]. The high contrast and thin index rings provide adequate effective index separation (>1.0×10^{-4}) between near eigenmodes and reduce the inter-mode crosstalk, which ensure the stable transmission of OAM modes. However, in turn, the mode effective area is reduced due to the high contrast and thin rings, resulting in unexpected high nonlinearity. Similar to conventional fiber communication systems, OAM transmission fibers are also susceptible to nonlinear effects, which will affect the fiber transmission capacity and lead to performance deterioration[8,9]. For OAM transmission, nonlinear effects can also cause nonlinearity-induced power loss and reduce the purity of OAM spectrum. Increasing the mode field area can effectively reduce the optical power density and the adverse effects caused by nonlinearity[10,11]. But on the other hand, the purity of OAM mode and the effective index separation between near eigenmodes will also be affected by the structural adjustment. Therefore, OAM transmission fibers with large mode area should be specially designed to take into account all these optical issues.

Multi-core supermode fibers[12-14], whose transmitted mode field energy is distributed in multiple cores, are easy to achieve large mode area, and have been used for high-power fiber transmission, high-power laser and space division multiplexing systems. The design and optimization of multi-core large mode area OAM transmission fibers need to adjust several interdependent and even contradictory optical properties at the same time, including mode area, nonlinear coefficient, effective index difference (Δn_{eff}), OAM purity, etc. Generally, the design and optimization of fiber structures mainly depend on various numerical simulation methods and manual trial-and-error framework. The traditional numerical simulation methods[15-17], such as plane wave expansion and finite element method, require a lot of computing resources, and the calculation and convergence process is very slow. The manual trial-and-error process depends heavily on the physical intuition and

* This work has been supported by the National Natural Science Foundation of China (Nos.92048301, 62020106004 and 11704283), in part by the Tianjin Municipal Education Commission (No.2018KJ146), and in part by the Opening Foundation of Tianjin Key Laboratory of Optoelectronic Detection Technology and Systems (No.2019LODTS004).

** E-mail: weihuang@tjut.edu.cn

design experience of researchers, which makes the optimization process time-consuming and inefficient. The relationships and laws between the optical properties of OAM transmission fibers are complex and obscure, which makes it difficult to obtain appropriate and desired structures only by intuitive experience and structural fine-tuning.

In recent years, neural networks have been proved to have excellent performance in the prediction of optical properties for complex optical structures[18-20]. The extremely fast computation speed of neural networks can be fully utilized and combined with various efficient optimization algorithms[21-23], providing a new approach for automatic optimization and reverse design. In this paper, neural network and different particle swarm optimization (PSO) algorithms are applied to realize the automatic optimization and design of large mode area multi-core OAM transmission fibers. The neural network has been established to achieve efficient and accurate predictions of multiple optical properties for different multi-core fibers, including mode area, nonlinear coefficient, purity, dispersion, and effective index difference. Then the trained neural network is combined with PSO and multi-objective particle swarm optimization (MOPSO) algorithms respectively to search for the desired fiber structures automatically. Due to the structural advantages of multi-core fiber and the automatic optimization process, we designed a number of multi-core structures with high OAM mode purity (>95%) and ultra-large mode area (>3 000 μm^2), which is larger by more than an order of magnitude compared to the conventional ring-core OAM transmission fibers.

In this section, multi-core fiber structure is applied to construct the structural framework of large mode area OAM transmission fibers, and the prediction neural network has been established to realize the fast and accurate prediction of the corresponding optical properties, including mode area, nonlinear coefficient, purity, dispersion, and effective index difference.

The prediction performance of neural network model depends on appropriate data sets and reliable training process. Before training the neural network, we took multi-core fiber structure as the design and optimization framework and established the data sets. The mode field energy in the multi-core fiber is distributed in all core channels, providing excellent condition for the design and optimization of large mode area OAM transmission fiber. As shown in Fig.1, the physical parameters of multi-core structure include the number of cores k, the core radius r, the distance L from each core to the fiber center, and the material refractive indices of each core and cladding n_1, n_2. These physical parameters are adjusted, and the corresponding optical properties are calculated and collected by COMSOL Multiphysics (commercial finite element code) to establish the data sets.

The adjustment range of each physical parameter in the data sets is set to $8\leq k\leq 10$, $0.02\leq(n_1-n_2)/n_2\leq 0.05$,

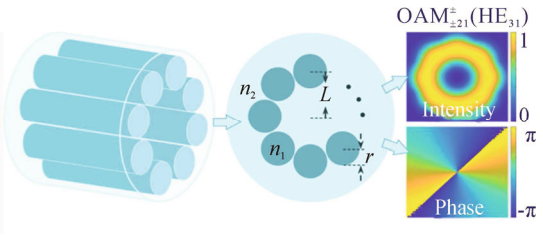


Fig.1 Cross section of the multi-core fiber structure

28 μm$\leq L\leq$30 μm, 15 μm$\leq r\leq$18 μm. As the OAM modes are formed by the linear combinations of conventional vector eigenmodes

$$\mathrm{OAM}^{\pm}_{\pm l,m}=\mathrm{HE}^{\mathrm{even}}_{l+1,m}\pm \mathrm{iHE}^{\mathrm{odd}}_{l+1,m},$$
$$\mathrm{OAM}^{\mp}_{\pm l,m}=\mathrm{EH}^{\mathrm{even}}_{l-1,m}\pm \mathrm{iEH}^{\mathrm{odd}}_{l-1,m}, \quad (1)$$

where l represents the topological charge number, the sign in superscript ± denotes the right or left circular polarization, and the sign of ±l denotes the right or left wave front rotation direction. Therefore, in this paper, 7 eigenmodes are selected to evaluate the optical performance of large mode area multi-core OAM fiber, including HE_{21}, TE_{01}, TM_{01}, EH_{11}, HE_{31}, EH_{21} and HE_{41} modes. The corresponding optical properties of the 7 eigenmodes calculated and collected in the data sets contain the mode field area (A_{eff}), nonlinear coefficient (γ), dispersion (D), effective index difference (Δn_{eff}) between adjacent higher order modes, and the purity (P) of the corresponding OAM modes. All the optical properties are calculated in the wavelength range of 1 500—1 600 nm with an interval of 3 nm. In total, 150 different structures are calculated and the optical properties of 7 eigenmodes at 34 wavelengths are collected to establish the data sets.

Before training the neural network model, 70% training samples are randomly selected from the 150 structures, 20% validation samples are randomly selected to provide an unbiased assessment while adjusting the parameters of the neural network, and the remaining 10% data samples are marked as test data set. After testing different parameters, the prediction neural network model is constructed with only one hidden layer. As shown in Fig.2, the calculated mean square error (*MSE*) values under different epochs for one-hidden-layer network are shown in Fig.2(a), and the *MSE* values for a two-hidden-layer network are shown in Fig.2(b). There is obvious over-fitting phenomenon in the figure, and the more hidden layers of the neural network, the earlier the over-fitting occurs in training process.

To address the overfitting issue, the neural network is constructed with only one hidden layer, and a dropout layer is added to the original one-hidden-layer neural network and the corresponding parameter is set to 0.1, i.e., each neuron has a 10% probability of being dropped. This will increase the sparsity of the neural network, making it easier to select features and prevent overfitting.

Due to the dropout operation, two neurons do not always appear in a dropout network, so that the update of weight values no longer depend on the joint action of hidden nodes with fixed relationship and the network can learn more robust features. In addition, L2 regularization methods are also added in the network, which will enforce certain constraints on the loss function during training process, so that the absolute value of the weight parameters will decrease with the update. As shown in Fig.3(a), the training loss and verification loss can converge to comparable values after the addition of the dropout layer and L2 regularization. We further increased the number of epochs to 3 100 and added the dynamic learning rate adjustment (the learning rate will be reduced to 10% of the original rate after every 1 000 epochs). The network can converge normally, and the oscillation amplitude decreases steadily after increasing the number of epochs, as shown in Fig.3(b).

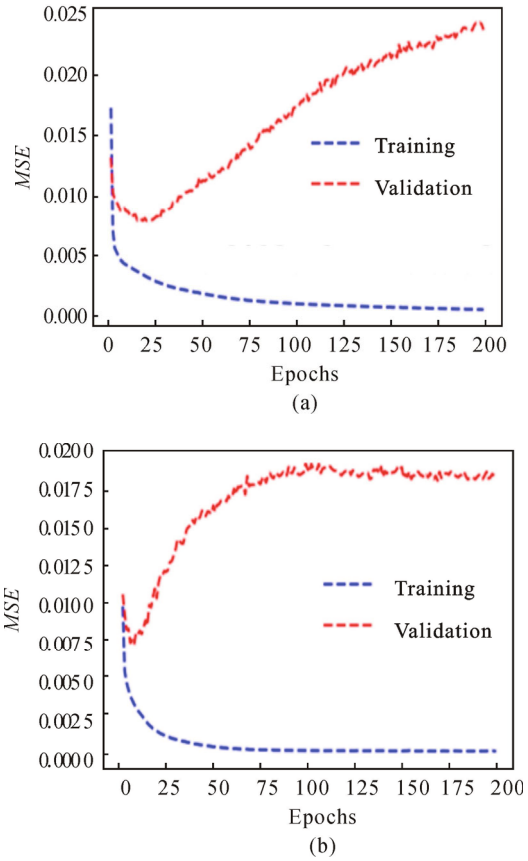


Fig.2 *MSE* values of the two networks under different epochs, each hidden layer of the two networks with 128 nodes: (a) Overfitting arises after about 25 iterations for one-layer network; (b) Overfitting arises after about 10 iterations for two-layer network

Next, we tested the performances of the prediction neural network models under different hidden layer nodes (32, 64 and 128 nodes). The number of hidden layer nodes has little effect on the final *MSE* convergence value. The network with 128 hidden layer nodes converges faster, so the number of hidden layer nodes of the prediction network is set to 128 finally. Additionally, similar to the number of hidden layer nodes, the batch-size affects the network convergence speed, and has little impact on the final *MSE* value. The batch-size is finally set to 32 for faster convergence. Adam optimizer is used to optimize the weight value during the training process.

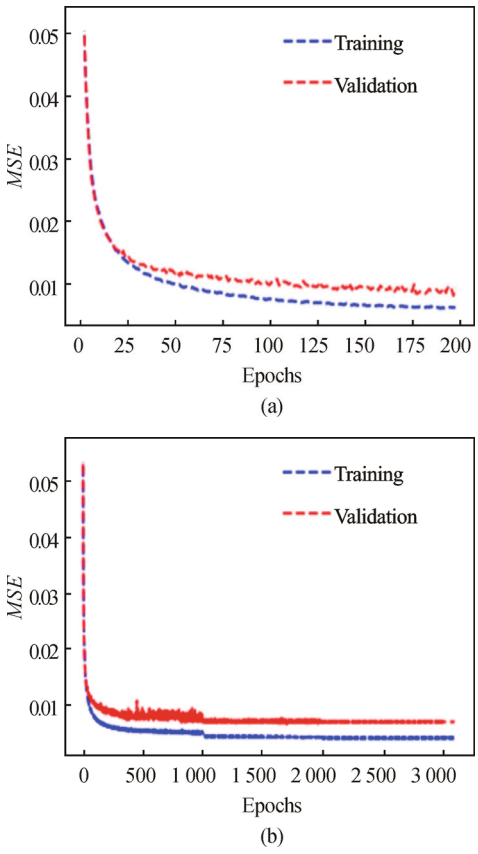


Fig.3 (a) The network can converge normally when the dropout layer and L2 regularization are added; (b) The network can converge normally, and the oscillation amplitude decreases steadily after increasing the number of epochs and changing the dynamic learning rate

Fig.4 shows the topology of the prediction neural network model, including an input layer, hidden layer, and output layer. The input layer has 5 nodes corresponding to the 4 structural parameters and wavelength λ. Here the parameter n is represented as $n=(n_1-n_2)/n_2$. The output layer is set as 54 nodes to represent the different optical properties of the selected vector modes: the mode field area (A_{eff}), nonlinear coefficient (γ), dispersion (D), and the purity (P) of the 12 even and odd vector modes (HE_{21A}, HE_{21B}, TE_{01}, TM_{01}, EH_{11A}, EH_{11B}, HE_{31A}, HE_{31B}, EH_{21A}, EH_{21B}, HE_{41A} and HE_{41B}), and the effective index difference (Δn_{eff}) between adjacent higher order modes (TE_{01}–HE_{11}, TM_{01}–HE_{11}, HE_{21}–HE_{11}, HE_{31}–HE_{21}, EH_{21}–EH_{11} and HE_{41}–HE_{31}).

Fig.5 shows the prediction error of each optical property, and the prediction error is defined as *Error* = (*predicted value* − *true value*)/*true value*×100%. The prediction errors

of different vector modes belonging to the same LP mode are similar, so in order to show the prediction results more clearly, we take HE_{21}, HE_{31} and HE_{41} modes to represent the OAM_{11}, OAM_{21} and OAM_{31} modes, respectively. All prediction errors are less than 5%. The prediction errors of mode purity, mode field area and dispersion are less than 3.5%, and 90% of the nonlinear prediction error data is less than 3.5%. The prediction results indicate that the neural network model can realize the fast and accurate prediction of optical properties for multi-core fiber structures.

After training and testing, the neural network model parameters are saved. To realize the optimization and reverse design of large mode area OAM transmission fiber, PSO and MOPSO algorithms have been applied and combined with the trained neural network respectively to search for the desired multi-core structures.

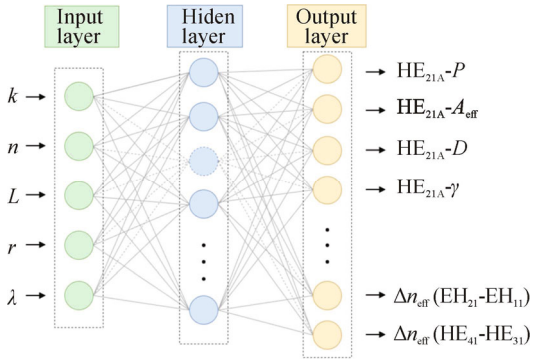
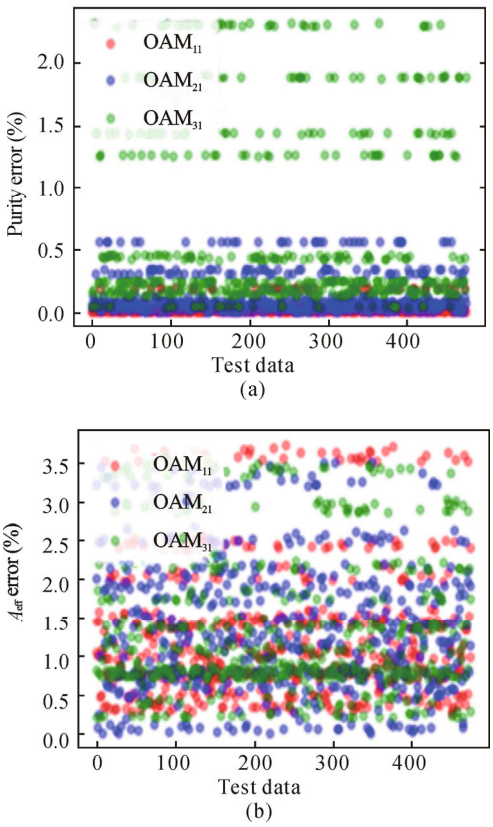


Fig.4 Topology of the prediction neural network model with 1 input layer (5 input nodes), 1 hidden layer, and 1 output layer (54 output nodes)



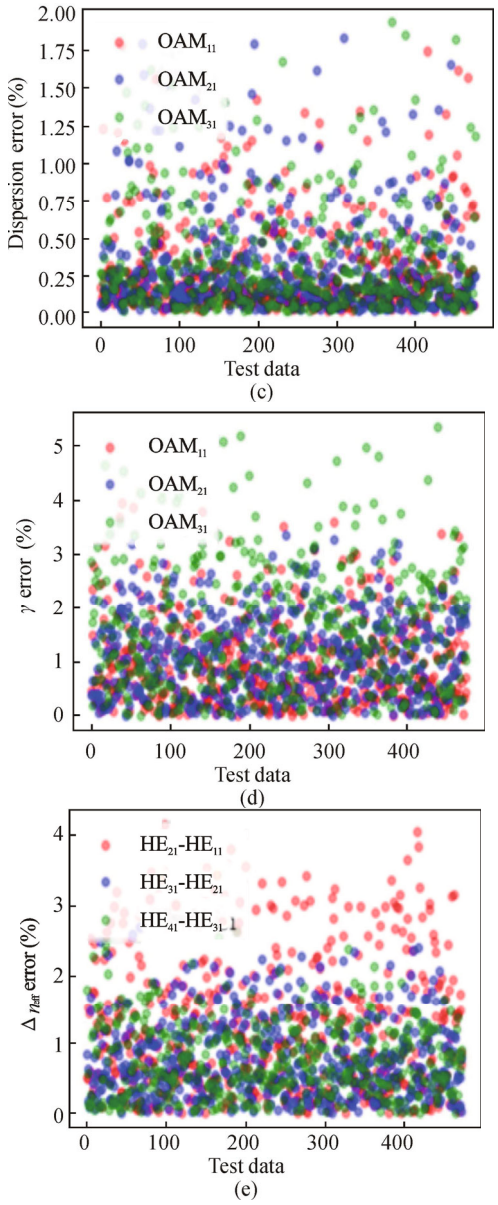


Fig.5 Prediction errors of (a) OAM purity, (b) mode field area, (c) dispersion and (d) nonlinear coefficient of OAM_{11} (represented by HE_{21}), OAM_{21} (represented by HE_{31}) and OAM_{31} (represented by HE_{41}) modes; (e) Prediction error of the effective index difference (Δn_{eff}) between adjacent higher order OAM modes (represented by HE_{21}-HE_{11}, HE_{31}-HE_{21}, and HE_{41}-HE_{31}, respectively)

The flow chart of PSO algorithm is shown in Fig.6. The particle swarm is initialized first and the parameters of particles are generated randomly within the specified range of $8 \le k \le 10$, $0.01 \le n \le 0.05$, $20\ \mu m \le L \le 30\ \mu m$, $10\ \mu m \le r \le 20\ \mu m$, $1\ 500\ nm \le \lambda \le 1\ 600\ nm$. The particle will be re-initialized randomly if the moving particle exceeds the defined boundary. Then the algorithm will calculate the defined fitness function value for each particle and update the personal best position (pbest), global best position (gbest) and moving velocity in every iteration. The calculation loop is executed until the maximum number of iterations is reached, and the algorithm will

output the optimal particle position and the optimal fitness function value.

The fitness function is defined according to different design requirements. For OAM transmission fibers, OAM purity is a very important optical property. In order to ensure the stability of OAM mode, we set the purity threshold to be greater than 95%. Fig.7 shows the calculation process based on neural network and PSO algorithm. Here, we only search for the multi-core structure with the largest mode field area (A_{eff}) in the PSO algorithm and the fitness function is defined as the sum of the mode field areas of the transmission modes:

$$fitness = \sum A_{\text{eff}}. \tag{2}$$

In the optimization process, the number of particles is set to 50, and the maximum iteration is set to 50 times. After 50 iterations, the algorithm will output a group of multi-core fiber structural parameters.

We have executed the PSO algorithm five times, and the corresponding five groups of output structures are shown in Tab.1. The fitness function values of the five structures are arranged in descending order.

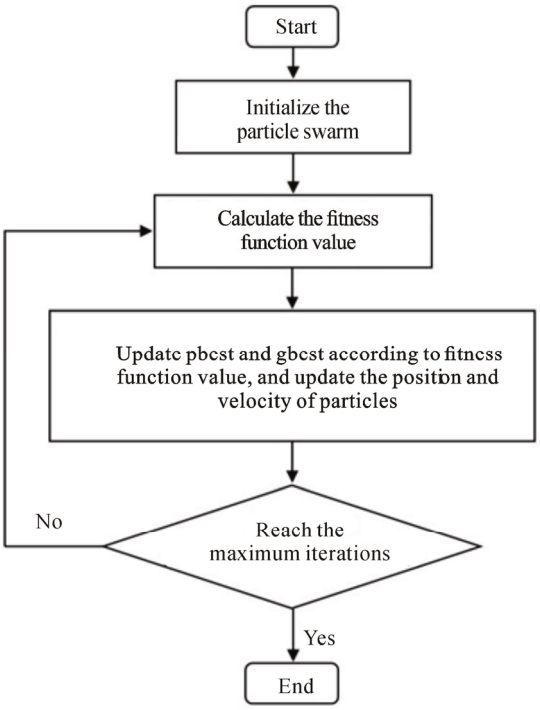


Fig.6 Flow chart of PSO algorithm

To verify the accuracy of the output structures by PSO algorithm, we select the first three groups of parameters in Tab.1 and calculated the corresponding optical properties by COMSOL Multiphysics. The corresponding transmission OAM purity of all the three groups of results exceeds 95%. The simulation results of mode areas are shown in Fig.8. The mode areas of all the selected 8 eigenmodes (including HE_{11} mode) exceed 4 000 μm^2,

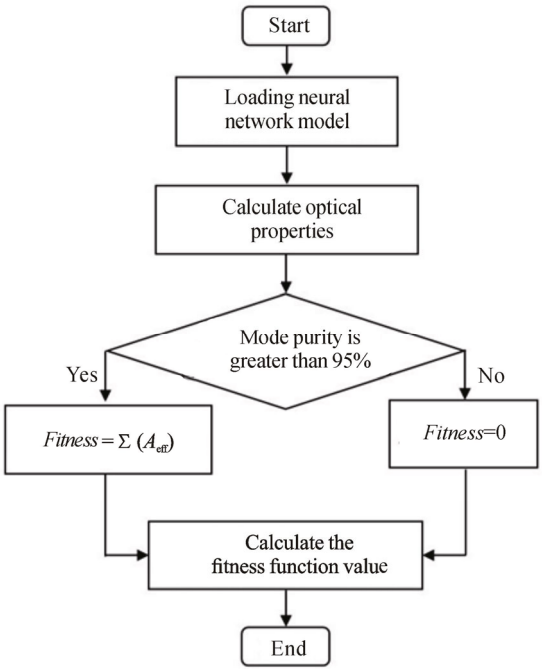


Fig.7 Calculation process based on neural network and the PSO algorithm

Tab.1 Five output results of the PSO algorithm

k	$n\times10^3$	L (μm)	r (μm)	λ (μm)	*Fitness*
10	18.37	29.905 4	19.755 1	1.548 57	50 470.8
10	14.21	29.816 1	18.108 5	1.534 79	50 401.7
10	11.25	29.796 3	17.999 8	1.553 44	50 393.8
10	15.18	29.727 3	18.197 8	1.548 18	50 337.9
10	26.60	29.856 9	17.025 6	1.569 37	49 334.3

and the largest mode field area exceeds about 4 800 μm^2, which is larger by more than an order of magnitude compared to the conventional ring-core OAM transmission fibers[4-7].

The PSO algorithm with only one fitness function provides a simple and direct approach to optimize and design multi-core fiber structures. But sometimes multiple optical properties need to be considered and optimized. For OAM transmission fibers, the effective index difference (Δn_{eff}) between adjacent higher order modes is also very important, and large Δn_{eff} can prevent mode coupling and reduce the crosstalk. Therefore, considering the effective index difference Δn_{eff}, MOPSO algorithm has been applied to optimize both mode area A_{eff} and Δn_{eff} for multi-core fibers. The calculation flow chart of MOPSO algorithm is shown in Fig.9. Two fitness functions are defined, and the sum of Δn_{eff} between four groups of adjacent higher order modes (HE_{21}–HE_{11}, HE_{31}–HE_{21}, EH_{21}–HE_{31} and HE_{41}–HE_{31}) is considered and defined as the second function.

$$fitness_1 = \sum A_{\text{eff}}, \tag{3}$$

$$fitness_2 = \sum \Delta n_{\text{eff}} \times 10^{10}. \tag{4}$$

Similarly, the number of particles in the MOPSO algorithm is set to 50, and the number of iterations is also set to 50. We have executed the MOPSO algorithm five times, and the corresponding five groups of output structures are shown in Tab.2. The corresponding transmission OAM purity of all the five groups of results exceeds 95%. The fitness function values of the five structures are arranged in descending order, and we also select the first three groups of parameters and calculated the corresponding optical properties (A_{eff}, Δn_{eff} and purity) by COMSOL Multiphysics to verify the design and optimization accuracy.

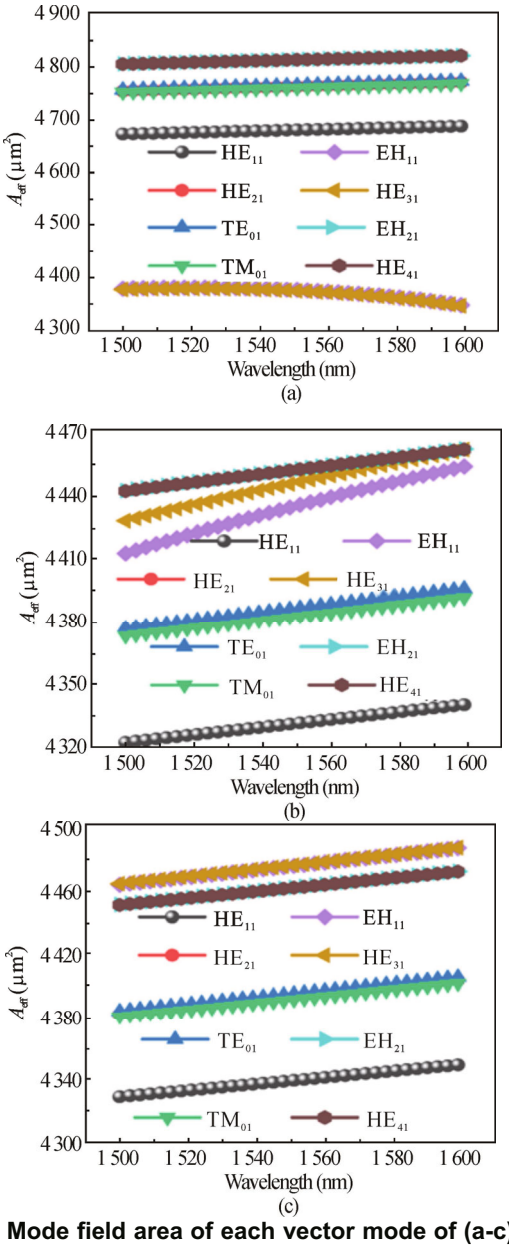


Fig.8 Mode field area of each vector mode of (a-c) the first three output structures in Tab.1 calculated by COMSOL Multiphysics

Fig.10 shows the refractive index differences between adjacent higher order modes of the first three output structures, and all the values of Δn_{eff} exceed 10^{-4}, which can prevent mode coupling and reduce the crosstalk in OAM transmission fibers. The mode areas of all the selected 8 eigenmodes (including HE_{11} mode) exceed 3 060 μm^2, as shown in Fig.11, which is also larger by more than an order of magnitude compared to the conventional ring-core OAM transmission fibers[4-7]. The corresponding OAM mode purity of the first three output structures is greater than 95%. The optimized output multi-core structures can support stable transmission of different OAM modes (OAM_{11}, OAM_{21} and OAM_{31}) and possesses ultra-large mode area, making it a good candidate for OAM communication system.

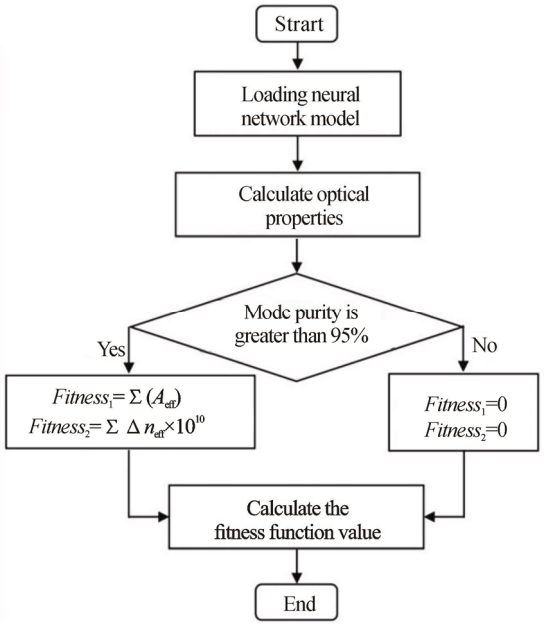
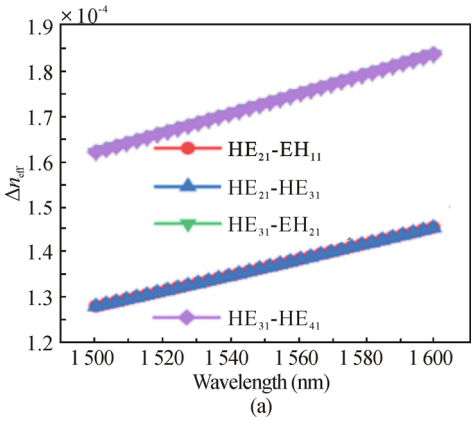


Fig.9 Calculation process based on neural network and the MOPSO algorithm

Tab.2 Five output results of the MOPSO algorithm

k	$n\times10^3$	L (μm)	r (μm)	λ (μm)	$Fitness_1$	$Fitness_2$
9	21.14	22.720 6	18.874 1	1.598 27	47 875.6	970 292.6
9	18.09	24.158 7	18.113 9	1.595 55	47 916.3	968 204.3
10	24.30	22.732 8	17.240 7	1.599 04	47 673.1	967 724.6
10	12.30	22.924 4	19.362 3	1.597 74	48 267.7	967 116.2
9	18.95	20.119 6	18.957 9	1.595 12	48 212.3	965 266.5



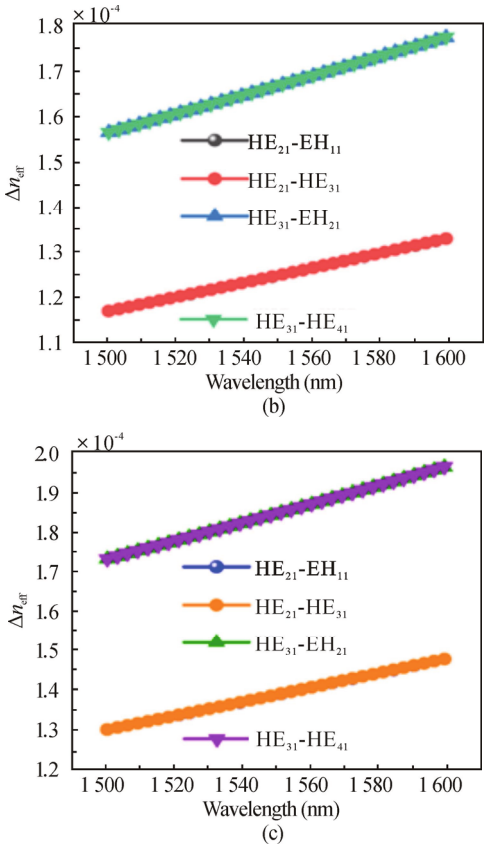
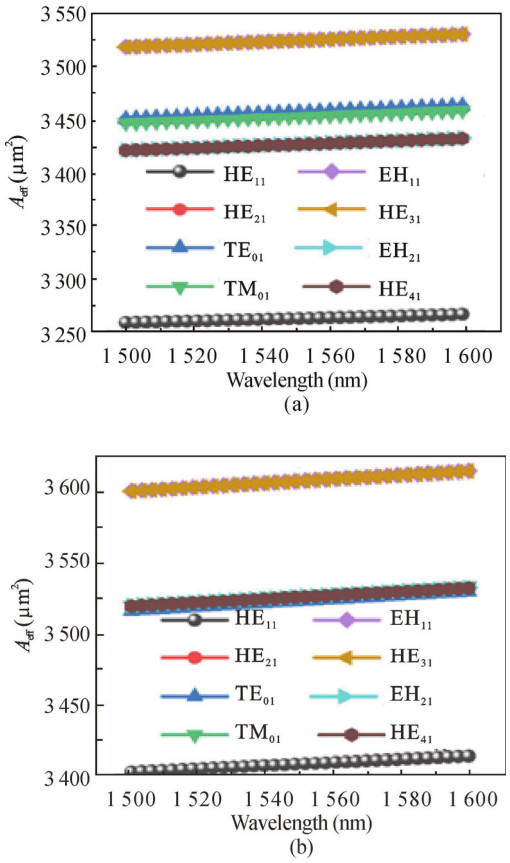


Fig.10 Refractive index differences between adjacent higher order modes of (a)-(c) the first three output structures in Tab.2 calculated by COMSOL Multiphysics



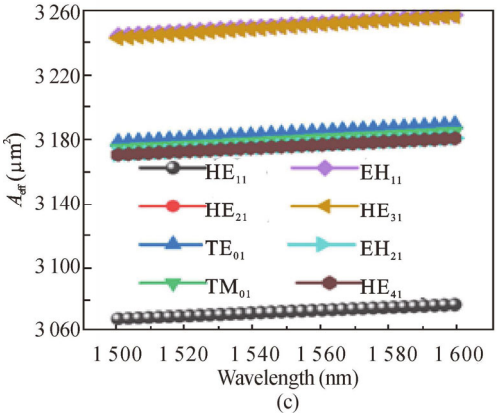


Fig.11 Mode field area of each vector mode of (a)-(c) the first three output structures in Tab.2 calculated by COMSOL Multiphysics

Ultra-large mode area multi-core OAM transmission fibers are proposed and optimized automatically based on accurate prediction neural network model and different PSO algorithms. Due to the structural advantages of multi-core fiber and the automatic optimization process, different multi-core structures with high OAM mode purity (>95%) and ultra-large mode area (>3 000 μm^2) have been proposed, which is larger by more than an order of magnitude compared to the conventional ring-core OAM transmission fibers.

Ethics declarations

Conflicts of interest

HUANG Wei and SONG Binbin are editorial board members for Optoelectronics Letters and were not involved in the editorial review or the decision to publish this article. All authors declare that there are no competing interests.

References

[1] MA Q C, LUO A P, HONG W Y. Numerical study of photonic crystal fiber supporting 180 orbital angular momentum modes with high mode quality and flat dispersion[J]. Journal of lightwave technology, 2021, 39(9)：2971-2979.

[2] WANG Q B, ZHANG X D, LIU X J, et al. Orbital angular momentum modes (de)multiplexer without mode conversion based on strongly guiding coupled ring-core fibers[J]. Journal of lightwave technology, 2022, 40(19)：6523-6533.

[3] FENG F, GAN J A, NONG J P, et al. Data transmission with up to 100 orbital angular momentum modes via commercial multi-mode fiber and parallel neural networks[J]. Optics express, 2022, 30(13)：23149-23162.

[4] TANDJÈ A, YAMMINE J, DOSSOU M, et al. Ring-core photonic crystal fiber for propagation of OAM modes[J]. Optics letters, 2019, 44(7)：1611-1614.

[5] QIN H B, HUANG W, SONG B B, et al. Hybrid method for inverse design of orbital angular momentum transmission

fiber based on neural network and optimization algorithms[J]. Journal of lightwave technology, 2022, 40(17)：5974-5985.

[6] ZHAO W Q, WANG Y N, LI S A, et al. Non-zero dispersion-shifted ring fiber for the orbital angular momentum mode[J]. Optics express, 2021, 29(16)：25428-25438.

[7] BANAWAN M, WANG L X, LAROCHELLE S, et al. Quantifying the coupling and degeneracy of OAM modes in high-index-contrast ring core fiber[J]. Journal of lightwave technology, 2021, 39(2)：600-611.

[8] SHEVCHENKO N A, NALLAPERUMA S, SAVORY S J. Maximizing the information throughput of ultra-wideband fiber-optic communication systems[J]. Optics express, 2022, 30(11)：19320-19331.

[9] MOELLER L. Nonlinear depolarization of light in optical communication fiber[J]. APL photonics, 2020, 5(5)：050801.

[10] VANVINCQ O, CASSEZ A, HABERT R, et al. Large mode area solid-core photonic bandgap Yb-doped fiber with hetero-structured cladding for compact high-power laser systems[J]. Journal of lightwave technology, 2021, 39(14)：4809-4813.

[11] ANUSZKIEWICZ A, FRANCZYK M, PYSZ D, et al. Nanostructured large mode area fiber for laser applications[J]. Journal of lightwave technology, 2022, 40(12)：3947-3953.

[12] LI S H, WANG J. Supermode fiber for orbital angular momentum (OAM) transmission[J]. Optics express, 2015, 23(14)：18736-18745.

[13] JOLLIVET C, MAFI A, FLAMM D, et al. Mode-resolved gain analysis and lasing in multi-supermode multi-core fiber laser[J]. Optics express, 2014, 22(24)：30377-30386.

[14] XIA C, BAI N, OZDUR I, et al. Supermodes for optical transmission[J]. Optics express, 2011, 19(17)：16653-16664.

[15] YU C P, CHANG H C. Applications of the finite difference mode solution method to photonic crystal structures[J]. Optical and quantum electronics, 2004, 36(1)：145-163.

[16] CUCINOTTA A, SELLERI S, VINCETTI L, et al. Holey fiber analysis through the finite-element method[J]. IEEE photonics technology letters, 2002, 14(11)：1530-1532.

[17] NORTON R A, SCHEICHL N. Planewave expansion methods for photonic crystal fibers[J]. Applied numerical mathematics, 2013, 63：88-104.

[18] GENTY G, SALMELA L, DUDLEY J M, et al. Machine learning and applications in ultrafast photonics[J]. Nanophotonics, 2021, 15：91-101.

[19] SINGH R, AGARWAL A, ANTHONY B W. Mapping the design space of photonic topological states via deep learning[J]. Optics express, 2020, 28(19)：27893-27902.

[20] WIECHA P R, ARBOUET A, GIRARD C, et al. Deep learning in nano-photonics：inverse design and beyond[J]. Photonics research, 2021, 9(5)：B182-B200.

[21] CHANG J H, CORSI A, RUSCH L A, et al. Design analysis of OAM fibers using particle swarm optimization algorithm[J]. Journal of lightwave technology, 2020, 38(4)：846-856.

[22] CHEN Y, GAO M, SONG X. Method to design the common aperture multi-band optical system based on the PSO algorithm[J]. Optics express, 2021, 29(12)：18325-18335.

[23] MOLESKY S, LIN Z, PIGGOTT A Y, et al. Inverse design in nanophotonics[J]. Nature photonics, 2018, 12：659-670.