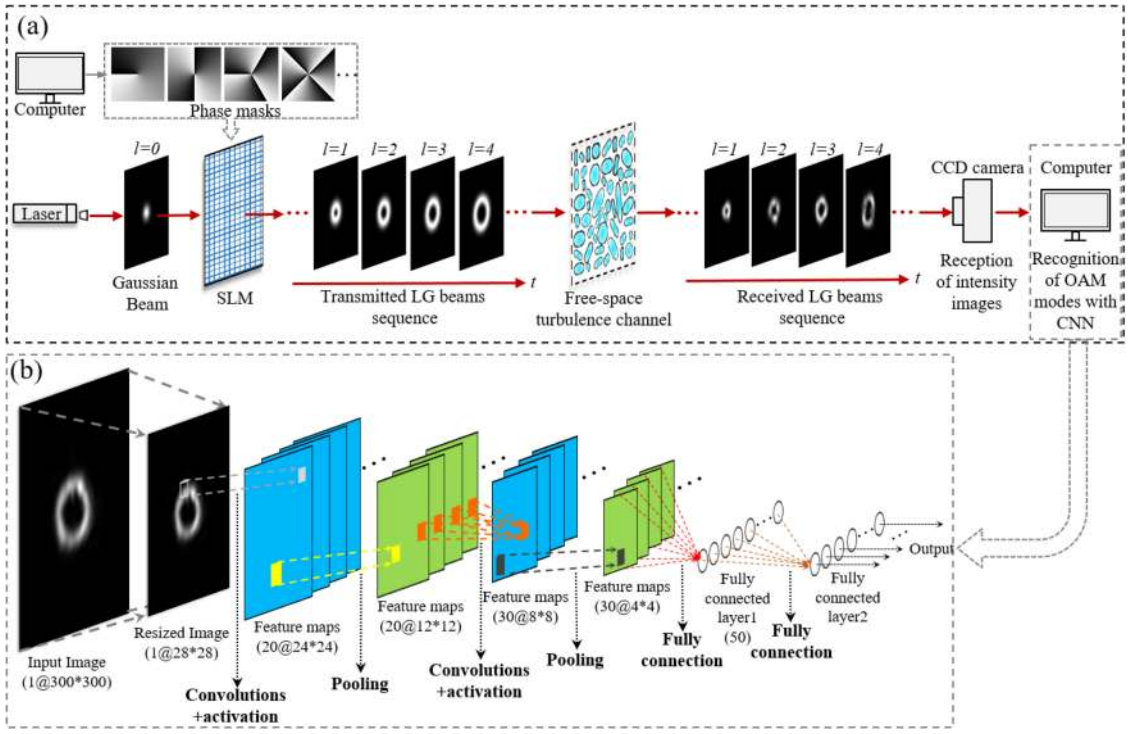
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Efficient Recognition of the Propagated Orbital Angular Momentum Modes in Turbulences With the Convolutional Neural Network

Volume 11, Number 3, June 2019

Zikun Wang
Maxime Irene Dedo
Kai Guo
Keya Zhou
Fei Shen
Yongxuan Sun
Shutian Liu
Zhongyi Guo



DOI: 10.1109/JPHOT.2019.2916207
1943-0655

Efficient Recognition of the Propagated Orbital Angular Momentum Modes in Turbulences With the Convolutional Neural Network

Zikun Wang,[1] Maxime Irene Dedo,[1] Kai Guo,[1] Keya Zhou,[2] Fei Shen,[1] Yongxuan Sun,[1] Shutian Liu,[2] and Zhongyi Guo [1]

[1]School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230009, China
[2]Department of Physics, Harbin Institute of Technology, Harbin 150001, China

DOI:10.1109/JPHOT.2019.2916207

Manuscript received April 8, 2019; accepted May 7, 2019. Date of publication May 10, 2019; date of current version May 24, 2019. This work was supported in part by the National Natural Science Foundation of China under Grants 61775050 and 61575060 and in part by the Fundamental Research Funds for the Central Universities JD2017JGPY0005, JZ2018HGBZ0309, and JZ2018HGTB0240. Corresponding author: Zhongyi Guo (e-mail: guozhongyi@hfut.edu.cn).

Abstract: The vortex beam carrying orbital angular momentum (OAM) has attracted great attentions in optical communication field, which can extend the channel capacity of communication system due to the orthogonality between different OAM modes. Generally, atmospheric turbulence can distort the helical phase fronts of OAM beams, which presents a critical challenge to the effective recognition of OAM modes. Recently, convolutional neural network (CNN), as a model of deep learning, has been widely applied to machine vision. In this paper, based on the CNN theory, we make a tradeoff between the computational complexity of the system and the efficiency of recognition by establishing a specially designed six-layer CNN structure in CPU station to efficiently achieve the recognition of OAM mode in turbulent environment through the feature extraction of the received Laguerre–Gaussian beam's intensity distributions. Furthermore, we examine the performances of our designed CNN with respect to various turbulence levels, transmission distances, mode spacings, and we have also compared the performances of recognizing single OAM mode with multiplexed OAM modes. The numerical simulation shows that basing on CNN method, the coaxial multiplexed OAM modes can obtain higher recognizing accuracy about 96.25% even under long transmission distance with strong turbulence. It is anticipated that the results might be helpful for future implementation of high-capacity OAM-based optical communication technology.

Index Terms: Pattern recognition, machine vision, optical vortices, free-space optical communication, atmospheric turbulence.

1. Introduction

To eventually overcome unsolved bandwidth crucial problem, realizing higher data transmission capacity is the main research direction of free-space optical (FSO) communication. Previous research fully exploited the wavelength [1], polarization [2], amplitude and phase [3], even spatially distribution of light fields [4]. Orbital angular momentum (OAM), a nature of Laguerre-Gaussian (LG)

beams with helical phase, was proposed and demonstrated firstly by Allen in 1992 [5]. It was shown that the LG beams comprising an azimuthal phase term of $\exp(il\varphi)$ have an OAM of $l\hbar$ per photon, where l (arbitrary integer) represents OAM mode called topological charge, φ is azimuthal angle, and $\hbar$ is Plank's constant h divided by 2π [5]. Benefited from the orthogonality between different OAM modes and the theoretically infinite range, the LG beams can be efficiently multiplexed and de-multiplexed to satisfy the ever-increasing capacity demands in FSO communication [6]–[11].

The effective recognition of OAM modes is the key to the OAM-based communication system. Typically, through spiral phase plate (SPP) with special spatial structure [10], grating [11]–[13], or dynamic loading varied computer-generated holograms (CGHs) corresponding with different OAM modes into spatial light modulators (SLMs) [14], incident beams carrying diverse OAM states can be demodulated into Gaussian-like beams and the initial information can be recovered by observing whether a bright spot is emerged on camera screen. Furthermore, a method based on the coordinate transformation was proposed to separate and recognize OAM modes [15]. However, the performances of the OAM-based FSO communication system expressly subject to the influence of the atmospheric turbulence (AT) which can randomly distort the phase front of a light beam and further induce channel crosstalk [16], [17]. To alleviate the AT effects, possible turbulence mitigation methods have been proposed, including aberration correction method where corresponding algorithm based on phase retrieval technique is used to correct the wave-front of distorted OAM modes [17], [18], and adaptive optics (AO) compensation method where a Gaussian beam is used to probe the turbulence-induced wave-front distortions [19].

Recently, machine learning technology has been widely applied to computer engineering as well as optical communication [20]–[23]. For OAM recognition, efficient artificial neural network (ANN) and K-nearest neighbor (KNN) based recognition methods were proposed, where intensity images of received beams were directly recognized and the corresponding OAM mode information could be successfully obtained [24]–[27]. However, in the process of processing the original image by ANN, a large number of domain experts are usually required to manually design the feature extractor to convert the form of the original image [27]. And KNN has problems of computational complexity and high sample imbalance. Comparatively, convolutional neural network (CNN) allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction, which is good at directly recognizing and discovering intrinsic features of input raw images and rarely occurs local minima or over-fitting [28]. The previous researches about recognizing OAM modes with CNN in OAM-based FSO communication were explored preliminarily in [27], [29]–[33]. The authors in [27] compared different machine learning algorithm models including CNN, ANN, KNN and naive Bayes classifier (NBC) in the multi-ary OAM transmission system. They further deepened the network hierarchy to achieve not only OAM demodulation but also atmospheric turbulence detection [29]. In Ref. [30], a trained CNN of Alexnet-architecture was used as a classifier to distinguish multiplexed Bessel-Gauss beams carrying OAM and tested it against traditional de-multiplexing method. The authors [31] designed an optical feedback network based on customized CNN with single layer of convolution and pooling, and demonstrated its ability to mitigate the effects of turbulence on OAM modes. They also investigated about using deep neural networks and CNN with single convolutional layer on classifying different OAM mode at different noise levels [32]. In Ref. [33], an OAM-SK-FSO communication system combining turbo coding method with a six-layer convolution and pooling layer of CNN was proposed.

In general, in these previous works, in order to achieve expected functions in specific cases, such as the correction of turbulence by the feedback network [31] and the exploration of noise levels [32], the single convolutional layer CNN with one convolution layer and one pooling layer was often used to minimize the complexity of the system. While, for the case of OAM mode recognition, CNNs with multi-hidden-layer were often used to optimize the recognition performance. For example, Ref. [29] used a seven-layer network structure including three layers of convolution and pooling layers and one layer of fully connected layer, and Ref. [30] used five-layer of convolutional layer, three-layer of pooling layer and three-layer of fully-connected layer with a total of eleven-layer of Alexnet-architecture model. And in Ref. [33], eight-layer of CNN was adopted to maximize the recognition accuracy. Although it has achieved good results in recognition accuracy, multi-layer networks require

more training parameters, longer computation time and also increases the complexity in practical operations. Besides, commonly used central processing unit (CPU) can no longer meet a large number of computational requirements, thus, graphics processing unit (GPU) capable of large-scale parallel computing were often used [30]. A trade-off between the computational complexity of the system and the efficiency of recognition was realized in [27], where a relatively moderate number of five-layer CNN structure consists of two convolution layer, two pooling layer and one fully connected layer were adopted, and the corresponding highest recognition accuracy of the 16-ary OAM was only around 84% after transmitting 2000 m under strong turbulence. Therefore, the better recognition performance after adopting a CNN with more suitable structure and optimization methods is what we need to pursue. Moreover, previous work did not give a quantitative analysis of the impact of multiplexed OAM modes with different mode spacings, and in most cases the multiplexed OAM modes were superposed by the symmetrical OAM modes with the same absolute value but opposite signs, because the formed petal-like intensity pattern was less insensitive to effects of the AT and relatively clearer for the CNN to distinguish [24]–[27]. So, the expansion of the number of the OAM modes and the multiplexed OAM modes based on different mode sets are still worth being investigated. In addition, the comparison of recognizing performance between the single OAM mode and the multiplexed OAM modes has also not been discussed before. These enlighten us to further explore and discuss the adopting of a specially designed CNN to recognize OAM modes on the performance of FSO communication system.

Different from above previous work, in this paper, we make a trade-off between the computational complexity of the system and the efficiency of recognition by establishing a specially designed six-layer CNN structure in CPU station to efficiently achieve the recognition of OAM mode through the feature extraction of the received LG beam's intensity distributions. We examine our designed CNN with respect to various turbulence levels, transmission distances and mode spacings and attempt to compare the performances of recognizing single OAM mode with that of multiplexed OAM modes. By comparison, we can obtain that coaxial multiplexed OAM modes can get higher recognizing accuracy even under long transmission distance with strong turbulence. The results may help to achieve high-capacity OAM-based optical communications technology in the future.

2. Concept and Principle

2.1 The Laguerre-Gaussian Beams Carrying Orbital Angular Momentum

There are different types of beam carrying OAM, one of them is the LG beam. For the cylindrical coordinates (r, φ, z), where r is the distance from the propagation axis, φ is the azimuthal angle, and z is the propagation distance, the field distribution of LG beam [34] can be expressed as:

$$\begin{aligned} u_{LG(l,p)}(r,\varphi,z) = & \sqrt{\frac{2p!}{(\pi(p+|l|)!)}}\frac{1}{w(z)}\left(\frac{r\sqrt{2}}{w(z)}\right)^{|l|} L_p^{|l|}\left(\frac{2r^2}{w^2(z)}\right) \\ & \times \exp\left(\frac{-r^2}{w^2(z)} - \frac{ikr^2 z}{2R(z)}\right) \\ & \times \exp\left[i(2p+|l|+1)\tan^{-1}\frac{z}{z_R}\right]\exp(il\varphi), \end{aligned} \tag{1}$$

where λ is the wavelength, $k = 2\pi/\lambda$ is the wavenumber, l and p are the mode order called topological charge and radial indices respectively, $w(z) = w_0\sqrt{1+(z/z_R)^2}$ is the Gaussian beam radius, w_0 is the beam waist, $z_R = \pi w_0^2/\lambda$ is the Rayleigh range, $L_p^{|l|}(\cdot)$ is the generalized Laguerre polynomial, and $(2p+|l|+1)\tan^{-1}(z/z_R)$ is the Gouy phase. It is important to note that the phase factor, expressed as $exp\,(il\varphi)$, is what allows these beams to exhibit OAM [34]. The intensity profiles and phase fronts of a LG beam with different OAM modes is shown in Fig. 1.

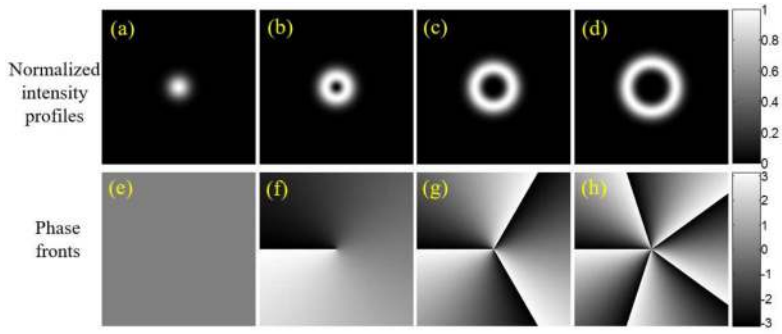


Fig. 1. (a)–(d) Simulated normalized intensity profiles and phase fronts of Laguerre-Gaussian (LG) beams with different OAM modes of $l = 0$ (Gaussian-like beam), $l = 1$, $l = 3$ and $l = 5$ respectively. (e)–(h) Phase fronts corresponding to (a)–(d) at $z = 0$.

2.2 The Analytical Model of Atmospheric Turbulence Channel

The atmospheric turbulence (AT) is caused by random variations in temperature and convective motion resulting in random variations in the air's refractive index, which can easily distort the phase front of a light beam [35]. Usually, phase fluctuations result in amplitude fluctuations as the beam propagates. In numerical simulation, turbulence channel was simulated by inserting random phase screens along the propagation path of the beam corresponding to the model developed by Hill [36] and defined analytically by Andrews [35]:

$$\phi_n(k_x, k_y) = 0.033C_n^2 \left[1 + 1.802\sqrt{\frac{k_x^2 + k_y^2}{k_l^2}} - 0.254\left(\frac{k_x^2 + k_y^2}{k_l^2}\right)^{\frac{7}{12}} \right] \times exp\left(1 - \frac{k_x^2 + k_y^2}{k_l^2}\right)\left(k_x^2 + k_y^2 + \frac{1}{L_0^2}\right)^{\frac{-11}{6}}, \quad (2)$$

where C_n^2 is the structure constant of the refractive index of air, which represents the strength of AT, k_x and k_y denotes the wavenumber in x and y direction respectively, L_0 is the outer scale of AT and $k_l = 3.3/l_0$, l_0 is equals to the inner scale of AT. The fluctuation of the wave-front phase is simulated through a random distribution with variance of σ^2. Then, the perturbation of the refractive index is approximately described by the Kolmogorov spectrum:

$$\sigma^2(k_x, k_y) = \left(\frac{2\pi}{N\Delta x}\right)^2 2\pi k^2 \Delta z \phi_n(k_x, k_y), \quad (3)$$

where N and Δx respectively denotes the size and the grid interval of the phase screen, $k = 2\pi/\lambda$ is the wavenumber and Δz is the interval between sequential phase screens. In Cartesian coordinate, the phase screen is expressed in the frequency domain through the fast Fourier transform:

$$\varphi(x, y) = FFT(C_{N\times N}\sigma(k_x, k_y)), \quad (4)$$

where $C_{N\times N}$ is a complex random array with a mean of 0 and a variance of 1, and subscript represents the distribution over a sampling grid of size $N \times N$. Fig. 2(a)–(b) shows two examples of turbulence phase screens with different C_n^2 valued in 1×10^{-15} $m^{-2/3}$ and 1×10^{-14} $m^{-2/3}$, and Fig. 2(c)–(d) shows the intensity images of received LG beam with OAM mode of $l = 5$ under the case of 1000 m transmission distance with turbulent environment corresponding to (a–b).

2.3 The Theoretical Model of Convolutional Neural Network (CNN)

Deep learning refers to a machine learning method with multiple levels of representation, obtained by composing simple but non-linear modules that each transform the representation at one level

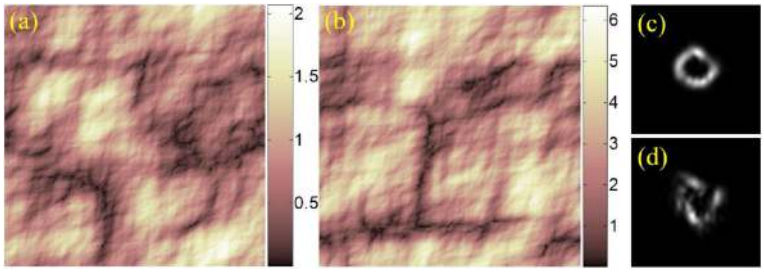


Fig. 2. (a), (b) Simulated turbulence phase screens with different C_n^2 valued in 1×10^{-15} $m^{-2/3}$ and 1×10^{-14} $m^{-2/3}$. (c), (d) The intensity images of received LG beam ($l = 5$) after transmitting 1000 m with turbulent environment corresponding to (a) and (b).

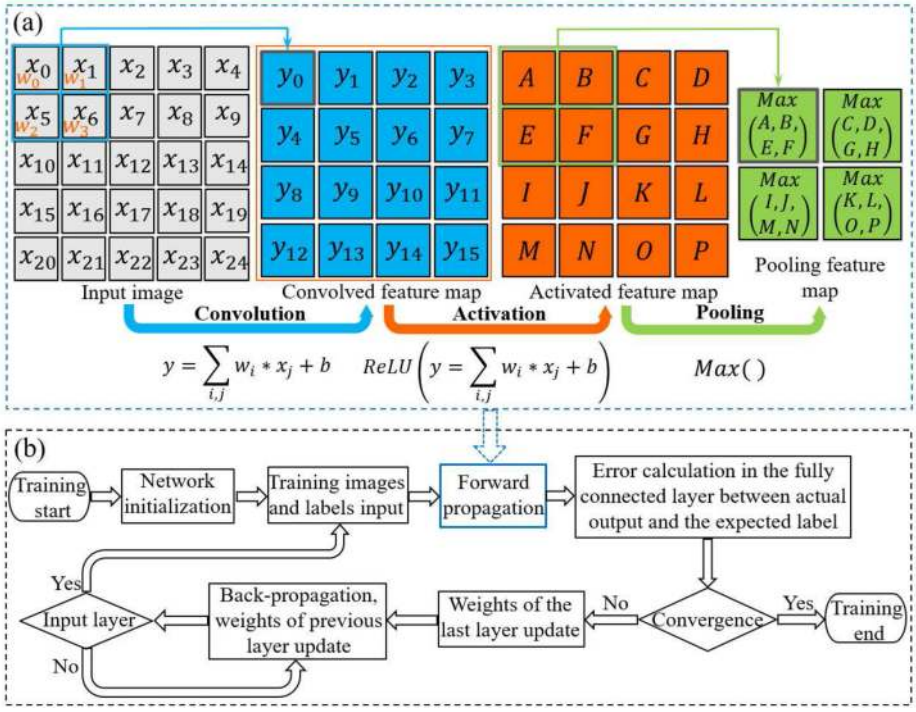


Fig. 3. The theoretical model of CNN. (a) Details of the convolution, activation and pooling operation in a single layer CNN. For example, the 5×5 input image is convolved with a 2×2 convolution filter and added a bias, then a 4×4 feature map is generated in the convolution layer. After being activated by a non-linear activation function such as a rectified linear unit (ReLU), the maximum in the non-overlapping 2×2 sub-region of the 4×4 feature map is calculated and then a 2×2 max-pooling feature map is emerged as an input of the next single layer. (b) The details of training process of CNN.

(starting with the raw input) into a representation at a higher, slightly more abstract level [28]. CNN, as a model of deep learning, is a trainable feed-forward multi-layer network structure composed of multiple single-layer convolutional networks. Benefiting from the advantages of CNN, such as local connection and weight sharing, it has been successfully applied in the fields of image and audio recognition.

Normally, each layer of CNN are composed of learned convolutional filters followed by a non-liner activation function and pooling operators. The details of the three operations of the single layer CNN are shown in Fig. 3(a). In the convolutional operation, with the designable spatial size of filters, stride, and padding, a collection of filter sets are trained and convolve with input images to extract specific

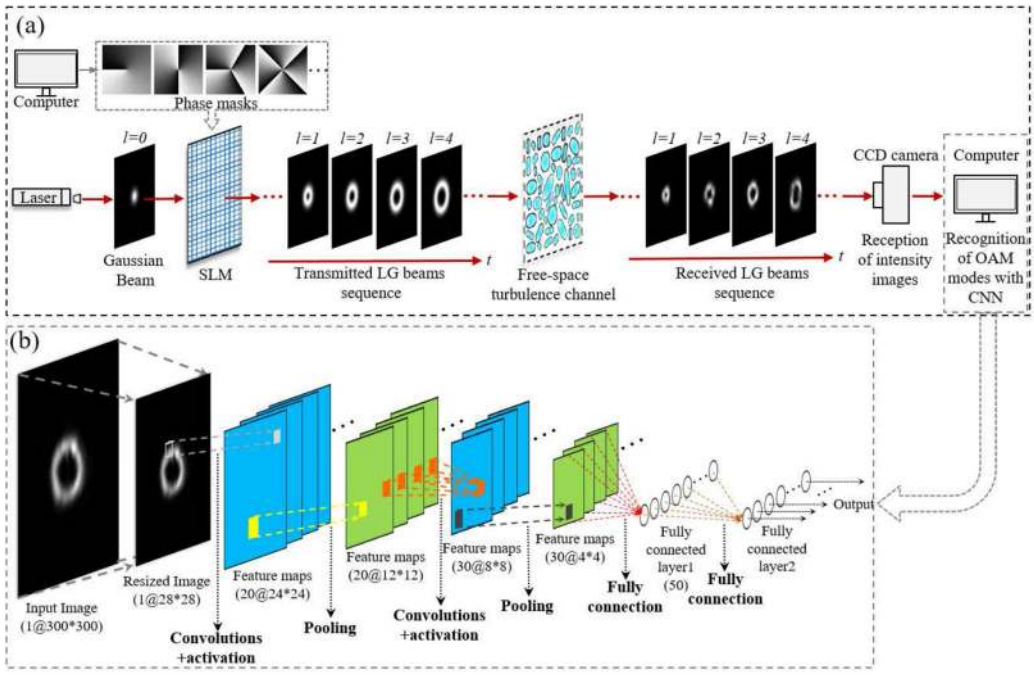


Fig. 4. Schematic diagram of OAM transmission system. (a) Structural details diagram of the OAM transmission system. (b) Structural details diagram of the OAM recognition method based on our designed CNN.

features at all locations, then generate the convolved feature maps. Due to the weight sharing, the activation of feature map can move synchronously with the shift of the input images, thus makes CNN less sensitive to the distortion of input image and further improves its generalization performance. After activated by a non-linear activation function to improve the non-linear expression ability of the network, the pooling operation takes inputs from a non-overlapping sub-region in the feature maps and extract the maximum or average of these inputs in order to dramatically reduce computational demands and add translational invariance. During the training period, after the training labeled images are forward propagated to multi layers of convolutional, activation and pooling operations, a series of features are extracted, and the notion of spatial information is abandoned, from which this multi-layer structure converts a level of representation from the original input to a higher, more abstract level of representation. And the subsequent operation is the fully connected operation. The fully connected layers consist of a non-linear activation function and a softmax classifier to achieve classification. When each labeled image has been processed by the network and forward propagated to the fully connected layers, a loss function measures the error, which is obtained by comparing the actual output feature map with the expected label. Then, this error is back-propagated through the network using the chain rule and the weights in every layer are updated using stochastic gradient descent. The details of the training process are shown in Fig. 3(b). Once the training is completed, input the test images, and the network in the testing period will perform classification operation, where the probabilities that the input belongs to every class, are respectively calculated, and the category corresponding to the maximum probability is further output by the CNN.

3. Numerical Simulation Results and Discussion

3.1 Simulation Description of the OAM Transmission System

The schematic diagram of the OAM transmission system under the AT influence is shown in Fig. 4. Fig. 4(a) shows the structural details diagram of the OAM transmission system. At the transmitter, we

simulate laser source to produce Gaussian beam ($l = 0$) with wavelength of $\lambda = 1550$ nm and beam waist of $w_0 = 3$ cm respectively, and then the beam is radiating onto phase masks loaded on spatial light modulator (SLM) that are programmed with different OAM modes for modulating Gaussian beams into corresponding LG beams. Next, the LG beams sequence are transmitted through the free-space turbulence channel, which is simulated by inserting random phase screens along the propagation path of the beam corresponding to the model developed by Hill and defined analytically by Andrews (that is described in detail in section 2.2). In the simulation, we have following model parameters: $N = 300$, $\Delta z = 200$ m, $L_0 = 50$ m, $l_0 = 0.0003$ m. And we adopt AT strength from weak turbulence to strong level, corresponding to C_n^2 values from 1×10^{-16} m$^{-2/3}$ to 1×10^{-14} m$^{-2/3}$.

After being affected by turbulence channel, at the receiver, the intensity images of the destroyed beams are captured by CCD camera and the method based on our specially designed CNN is adopted to recognize the OAM of the received LG beams sequence. It is worth mentioning that in the field of computer engineering, the recognition accuracy is often improved by deepening the number of network layers, but this will bring a lot of training parameters, lead to too long training time and introduce the problem of over-fitting. On the contrary, CNN with fewer number of layers (such as a CNN with single convolutional layer) is relatively easy to design, and it does reduce computational complexity as well as training time. However, the features extracted from too few layers are relatively fewer, which brings about a problem that the recognition accuracy is relatively low. Therefore, how to make a trade-off between the computational complexity and recognition accuracy (that is to say, rational design of the number of layers or the adoption of appropriate optimization methods) is the key to our designed CNN.

In our simulation, in order to reduce the amount of training parameters and the computation time than previous works [29]–[31], and meanwhile, guarantee to extract sufficient features and achieve a higher recognition accuracy than previous work [27], [29] under long transmission distance, we make a trade-off between the computational complexity of the system and the efficiency of recognition by establishing a specially designed six-layer of CNN in which each of convolution layer, pooling layer and fully connected layer have two layers, and various optimization methods have been adopted. It should be pointed out that the moderate number of layers are adopted here, which can not only ensure a better recognition performance by extracting more features than the CNN structure with fewer layers [27], [31], [32], but also avoid huge computation parameters or over-fitting problem that may cause by excessive number of layers [29], [30], [33]. At the same time, due to its smaller computation costs, this network can be established in local CPU station. The structural details are shown in Fig. 4(b). In the input part, the pixels of each received intensity images of LG beams are resized from 300×300 to 28×28 in order to reduce training parameters and improve training speed. Next is the convolution operation, where filters are convolved with the input to create a number of convolutional outputs and activated feature maps with 24×24 pixels, dependent on the filter number 20, pixel size 5×5, padding 0, and stride 1. In the pooling operation, 20 feature maps with 12×12 pixels are generated as inputs of the next layer through calculating the maximum in the non-overlapping 2×2 sub-region (stride 2) of the feature map with 24×24 pixels to reduce data size and improve the spatial translational invariance (distortion and displacement robustness). After performing similar operations in the later, all feature maps are connected to the fully connected layer with 50 nodes, and next, the second fully connected layer is connected with the previous 50 nodes. Then a softmax classifier is followed to achieve classification and get the output results.

Different from the past works, we have further adopted various optimization methods together during the training period to improve the performance of this specially designed CNN. Firstly, the rectified linear unit (ReLU) is chosen as the non-linear activation function in the convolution, pooling and fully connected layers due to its faster convergence rate than the traditional non-linear function such as sigmoid and hyperbolic tangent (Tanh) function [37]. Then, in order to better improve the stability of convergence, mini-batch method which equally divides the input training databases into several batches is used. In this case, training images in a batch traverse the whole network once and the errors are back-propagated with the weights updating during an iteration. Then in an epoch, the above operations are repeated for all batches, that is, all batches which consist of whole training databases traverse the whole network during one epoch of all epochs. Besides, the model relies on

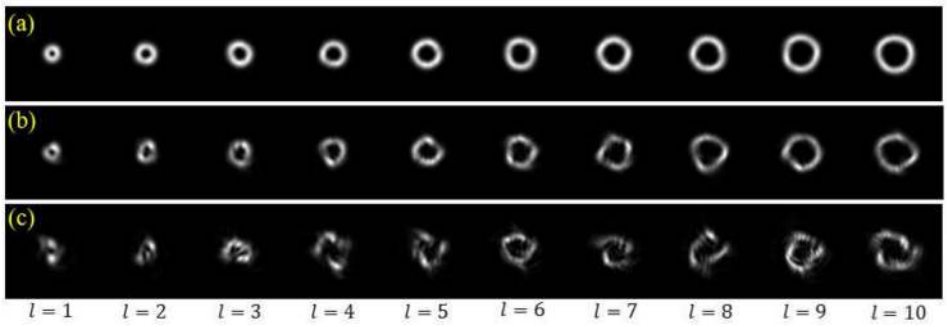


Fig. 5. The intensity images of received LG beam ($l = 1, 2, 3, ..., 10$) after transmitting 1000 m with AT strength levels $C_n^2 = 1 \times 10^{-16}$ m$^{-2/3}$, $C_n^2 = 1 \times 10^{-15}$ m$^{-2/3}$ and $C_n^2 = 1 \times 10^{-14}$ m$^{-2/3}$ corresponding to (a), (b) and (c).

gradient descent for training, where the learning rate is an important factor affecting convergence speed of the gradient descent and recognition accuracy. Thus, we use a momentum-based gradient descent method, namely, the learning rate decreases exponentially as the training epoch increased, which can accelerate the training speed of the model while improving the convergence ability. At last, we add a dropout unit between the last two layers of the fully connected layer of the designed CNN. The function of this unit is to stop the activation of a certain neuron with a certain probability and the neurons which are dropped out in this way do not contribute to the forward-propagation and do not participate in back-propagation. So inputting an image every time, the network samples a different structure, which reduces complex co-adaptations of neurons, since a neuron cannot rely on the presence of particular other neurons [37]. Therefore, it force our designed CNN to learn more robust features, make the CNN more generalizable and then avoids over-fitting.

3.2 Performance of the OAM Recognition Method Based on CNN

In this section, we analyze the performance of the OAM recognition method based on our designed CNN. To demonstrate the destructive effect of turbulent channels on LG beams, we simulate 10 single OAM modes ($l = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10$) after transmitting 1000 m under three AT strength levels corresponding to weak, medium and strong turbulence, as shown in Fig. 5. We can see that, as the AT is strengthened, wave-front phases will be damaged gradually, thus the intensity profile of diverse LG beams cannot be maintained, which makes it more challenging to recognize them accurately.

In order to reduce the impact of AT on the recognition of OAM modes, recognition method based on CNN is adopted at the receiver. Here, we define recognition accuracy to measure the performance of system:

$$Accuracy = \frac{\sum_{n=1}^{N} f(n)}{N} \times 100\%, \tag{5}$$

and

$$f(n) = \begin{cases} 1 & l_n = l_n^* \\ 0 & l_n \neq l_n^* \end{cases}, \tag{6}$$

where l_n is the training true label of LG beam, l_n^* represents the classification label predicted by CNN and N is the sample number. The performances are shown in Fig. 6. For transmitting single OAM mode, 10 single OAM modes are transmitted through AT respectively and recognized by CNN at the receiver. Firstly, we investigate the effect of different AT strengths on recognition accuracy, in which each of AT strength is divided into ten levels, for each level, 2500 intensity images of LG beams are accumulated and divided into training set and testing set (2000 for training and 500 for testing set). And the CNN training parameters are as follows: epoch = 20, batch-size = 100,

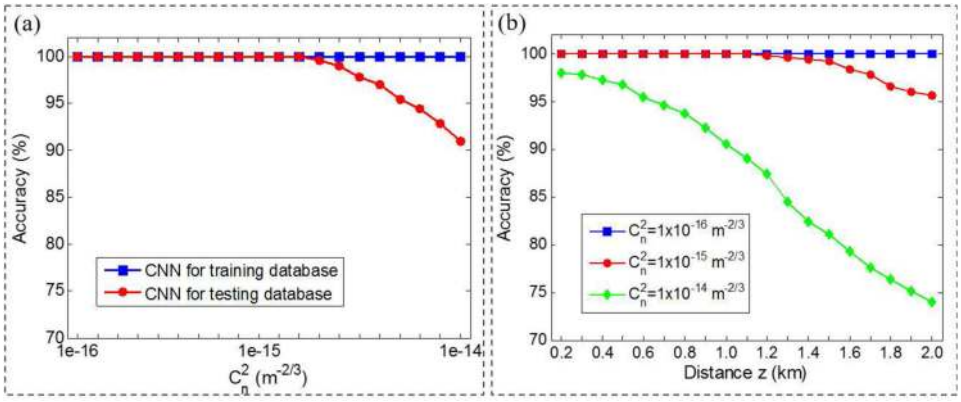


Fig. 6. The performances of recognizing 10 single OAM modes ($l = 1, 2, 3, ..., 10$) based on CNN method. (a) Accuracy of recognizing 10 single OAM modes with 1000 m transmission distance under different AT strength. (b) Accuracy of recognizing 10 single OAM modes under three AT strength levels with different transmission distance.

momentum = 0.5, dropout probability = 0.5. As we can see in Fig. 6(a), for the training database, the total training time of all levels is approximately 4.4 minutes, and the accuracy is 100% for all of the situations, which represents that as the number of iterations increases, finally, features of input images can be extracted and trained effectively. For the testing database, when C_n^2 is less than 1×10^{-15} $m^{-2/3}$, medium turbulence level, the recognizing accuracy maintains at 100%, but it declines to 91% when C_n^2 increases from 1×10^{-15} $m^{-2/3}$ to 1×10^{-14} $m^{-2/3}$, in which the AT strength increase from medium turbulence level to strong turbulence level. We can see that the CNN can effectively classify the OAM modes of the transmitted LG beams after the AT, and maintain the accuracy of more than 90% even in the case of strong turbulence strength levels.

Then we test the relationship between the accuracy of the OAM recognition and the transmission distance. We set three different levels of C_n^2 in simulation as shown in Fig. 5 and compare the accuracy versus different transmission distance z. The total training time of each levels is approximately 4.5 minutes and the results are shown Fig. 6(b), for $C_n^2 = 1 \times 10^{-16}$ $m^{-2/3}$ corresponding to the weak AT strength, no matter how far the transmission distance is (even the transmission reaches 2000 m), the accuracy remains at 100%. And when the AT strength is medium level with $C_n^2 = 1 \times 10^{-15}$ $m^{-2/3}$, the accuracy is still greater than 95% with 2000 m transmission distance. However, under strong turbulent environment, the accuracy is above 95% only within a short transmission distance of less than 500 m. With the continuous increase of transmission distance, the accuracy decreases rapidly, and when the transmission distance reaches 2000 m, it drops to only about 74%. Therefore, the accuracy of the OAM recognition method based on our designed CNN can achieve 100% accuracy within 1000 m in the case of weak to medium atmospheric turbulence level, but cannot adapt to strong turbulence level while maintaining long distance transmission.

It should be noted that when the LG beam carrying OAM propagates through the free space, the beam is constantly diverging. And due to the AT influence, the wave-front is severely distorted and the divergence becomes random and chaotic, resulting in OAM spectrum expansion which called channel crosstalk in communication filed. Normally, crosstalk between different OAM modes varies according to the mode spacing, therefore, we will also investigate the impact of different mode spacings on the accuracy of CNN recognition. Table 1 describes the three mode sets used for recognizing single OAM mode in simulation. Each set consists of 5 OAM modes $\{1, 2, 3, 4, 5\}$, $\{1, 3, 5, 7, 9\}$ and $\{1, 4, 7, 10, 13\}$ with different spacing of $\Delta = 1$, 2 and 3 between two adjacent modes. Based on these, we explore the accuracy of CNN in recognizing single OAM mode under the effect of long-distance (2000 m) turbulent environment. For each set, the average training time was 7 seconds at a particular turbulence intensity level. The results are presented in Table 2, as we can see, under the weak turbulence intensity, the choice of mode set has no effect on the

TABLE 1
Mode Set Used in Simulation

Set	For single OAM
1	$\{1, 2, 3, 4, 5\}$
2	$\{1, 3, 5, 7, 9\}$
3	$\{1, 4, 7, 10, 13\}$

TABLE 2
The Performance of Recognition Accuracy Based on Different Sets for Single OAM Mode

Results \ Sets	$C_n^2 = 1 \times 10^{-16} m^{-2/3}$			$C_n^2 = 1 \times 10^{-15} m^{-2/3}$			$C_n^2 = 1 \times 10^{-14} m^{-2/3}$		
	Set1	Set2	Set3	Set1	Set2	Set3	Set1	Set2	Set3
Accuracy (%)	100.00	100.00	100.00	98.00	98.80	100.00	82.00	91.60	96.00

recognition accuracy of CNN, and all sets can be fully recognized with accuracy of 100%. However, when the turbulence intensity increases to medium and strong levels with C_n^2 as 1×10^{-15} $m^{-2/3}$ and 1×10^{-14} $m^{-2/3}$ respectively, the recognition accuracy of set 3 with lager mode spacing is the highest and the set 1 with smaller mode spacing is the lowest. The reason is that, at the transmitter, for the single OAM mode, the intensity images corresponding to each OAM mode are all ring-shaped "doughnut" shapes and only the size of the ring is slightly different (as shown in Fig. 5). When LG beams with small mode spacing are transmitted into free-space turbulence channel, the received feature changes of intensity images corresponding to different OAM modes are mutually similar, thus classification errors will occur especially when CNN is classifying images corresponding to adjacent single OAM mode at the receiver. On the contrary, when transmitting LG beams with larger OAM mode spacing, it is easier to be classified and recognized because of the obvious feature difference in ring size, so the recognition accuracy is relatively high. But it is worth noting that, the mode spacing cannot be set too large, because increasing the spacing requires the inclusion of modes with higher mode order. Previous study has shown that channel efficiency decreases with relation to mode order [16], and LG beam with higher mode order has a wider beam radius and a lager contact area with turbulence filed, which means the beams would be degraded by turbulence filed more severely, leading to more crosstalks.

Above we consider the case of transmitting the single OAM mode, and in order to represent more bits and increase the capacity of the system, the multiplexing LG beam with m different OAM modes is usually adopted in the actual optical communication system because coaxially propagating LG beams with different OAM modes are mutually orthogonal, which is called the OAM multiplexing technology. In theory, the number m can be any positive integer, thus, infinite superposition can be chosen to configure the encoding mode set. It is worth mentioning that, in most past works where the multiplexed OAM mode was recognized, the multiplexing OAM modes with the same absolute value but opposite signs were often chosen in the simulations and experiments, because the petal-like intensity pattern formed by this superposition are less insensitive to effects of the AT and relatively clearer for the CNN to distinguish [24]–[27]. Here, we expand the number of superimposed OAM modes and explore the performance of recognizing multiplexed OAM mode based on our designed CNN. 4 ($m = 4$) OAM mode set $\{l = -4, -1, 2, 5\}$ is used as an example and multiplexed by complex phase masks loaded on SLM to create $2^m = 2^4 = 16$ different coherent superposition. The parameters of CNN are the same as those described above for recognizing single OAM mode.

The performances of recognizing multiplexed OAM modes are shown in Fig. 7. Fig. 7(a) shows the recognition accuracy of 16 different multiplexed LG beams under different AT strengths, as we can see, for the training database, input images' features are extracted and trained completely

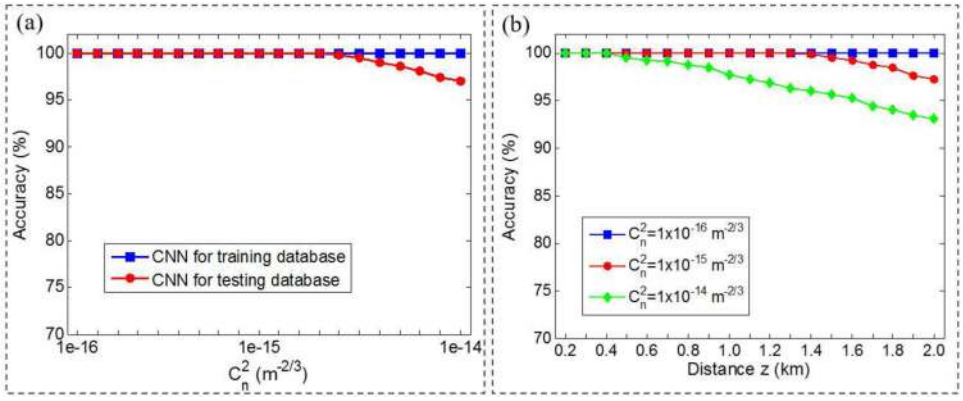


Fig. 7. The performances of recognizing multiplexed OAM modes (16 different superposition carrying 4 OAM modes $\{l = -4, -1, 2, 5\}$ based on CNN method. (a) Accuracy of recognizing multiplexed OAM modes with 1000 m transmission distance under different AT strength. (b) Accuracy of recognizing multiplexed OAM modes under three AT strength levels with different transmission distance.

TABLE 3
Mode Set Used in Simulation

Set	For multiplexed OAM
1	$\{-4, -1, 2, 5\}$
2	$\{-5, -1, 3, 7\}$
3	$\{-7, -2, 3, 8\}$

with the total training time of about 7.2 minutes. And for the testing database, the recognizing accuracy maintains at 100% originally, but declines slightly to 97% when turbulence level increases to strong turbulence level with $C_n^2 = 1 \times 10^{-14}$ $m^{-2/3}$. Then, we explore the relationship between the recognition accuracy and the transmission effect with the total training time of about 7 minutes and the test results are shown in Fig. 7(b). With the increase of transmission distance, when the transmission turbulent environment of free-space is weak level with C_n^2 valued in 1×10^{-16} $m^{-2/3}$, the accuracy curve is basically kept at 100%, and drops to about 97% when it is transmitted for 2000 m under the medium level turbulent environment with C_n^2 valued in 1×10^{-15} $m^{-2/3}$. Although in the case of strong level turbulence, the accuracy curve starts to decline gradually from about 500 m, and drops to about 93% after transmitting 2000 m, it is still higher than the accuracy of about 84% in Ref. [27] and about 90% in Ref. [29] under the same transmission distance with the same turbulence environment.

In addition, the impact of different mode spacings on the accuracy of CNN recognition has also been explored. Table 3 describes the three mode sets used for recognize multiplexed OAM mode under 2000 m transmission distance in simulation. Each set consists of 4 modes and all approximately centers around $l = 0$ with different spacings between adjacent modes. The intensity images of 16 coherent superposition corresponding to each mode set are shown in Fig. 8 and the performance of recognition accuracy for multiplexed OAM sets are presented in Table 4, where the average training time was 22 seconds for each set at a particular turbulence intensity level. We can see that, under the weak turbulence intensity, the choice of mode set has no effect on the recognition accuracy of CNN, in which all sets can be fully recognized. However, when the turbulence intensity level increases, the recognition accuracies of different mode sets are roughly similar, that is, the mode spacing seems to has little effect on the OAM multiplexing system. The exact reason for this condition should be investigated further in the future, but we preliminarily assume that the reason may be attributed as that for each set, the light intensity distributions of the 16 coherent

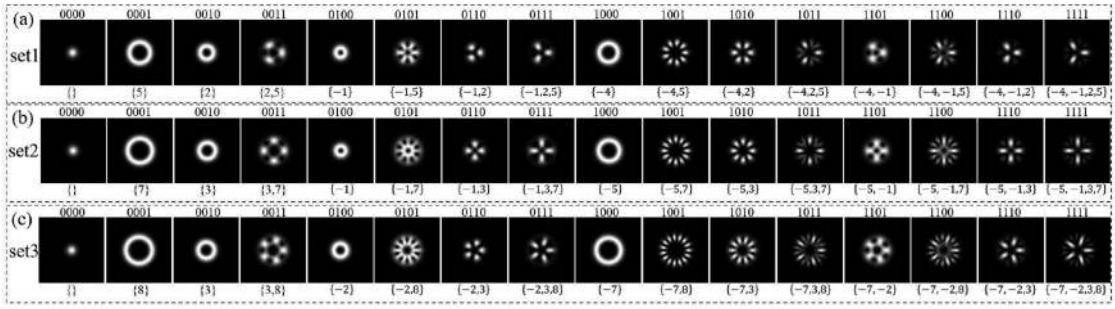


Fig. 8. Different coherent multiplexed LG beams based on three different OAM mode sets $\{l = -4, -1, 2, 5\}$, $\{l = -5, -1, 3, 7\}$ and $\{l = -7, -2, 3, 8\}$. The title above the sub-image is the encoded bit string (4 digit binary number) and the label below is the corresponding OAM modes.

TABLE 4
The Performance of Recognition Accuracy Based on Different Sets for Multiplexed OAM Modes

Sets / Results	$C_n^2 = 1 \times 10^{-16} m^{-2/3}$			$C_n^2 = 1 \times 10^{-15} m^{-2/3}$			$C_n^2 = 1 \times 10^{-14} m^{-2/3}$		
	Set1	Set2	Set3	Set1	Set2	Set3	Set1	Set2	Set3
Accuracy (%)	100.00	100.00	100.00	98.25	99.50	98.37	95.00	96.25	96.12

superposition are completely different from each other, as shown in Fig. 8, in which the features are obvious and each superposition has its own unique intensity distribution. Therefore, although the distribution of light intensity is destroyed after strong turbulence, due to the large difference in features between each intensity image, the probability of error is small when CNN classifies it. That is why the test recognition accuracy of the sets with different mode spacing is not much different in Table 4.

3.3 Discussion

Here, we analyze the different performances of recognition accuracy between multiplexed OAM modes as shown in Fig. 7 and single OAM mode as shown in Fig. 6. By comparing Fig. 7(a) with Fig. 6(a), we find that, under different AT strengths, the recognition accuracy curve of multiplexed OAM beams is slightly higher than that of single OAM beam, and the former keeps nearly 100% accuracy before turbulence level with $C_n^2 = 5 \times 10^{-15}$ m$^{-2/3}$, while the latter is $C_n^2 = 3 \times 10^{-15}$ m$^{-2/3}$. That is to say, under turbulent environment, the recognition performance of the multiplexed OAM beam is better than that of the single OAM beam. The reason has been mentioned above, and compared to all ring-shaped feature of single OAM mode as Fig. 5 shows, the features of the intensity images corresponding to different superposition modes are significantly different as shown in Fig. 8, thus the CNN can fully extract these features and perform high recognition accuracy. Similar result can be found by comparing Fig. 7(b) with Fig. 6(b), as the transmission distance changes, the accuracy curve of the multiplexed OAM modes is always higher than or equal to the accuracy curve of the single OAM mode at the same transmission distance. Even when transmitted long-distance with 2000 m under strong turbulent environment, the accuracy of the former was nearly 20% higher than the latter. At the same time, by comparing Table 4 with Table 2, for different mode spacings, the recognition rate of the multiplexed mode is substantially higher than the recognition rate of the single mode, and this is even achieved when the former has 16 classification and the latter only has 5 classification, which is also because the feature map of the former is more obvious than the later. These results indicate that due to good fidelity and recognition, the performance of using CNN to recognize multiplexed OAM modes is better than recognizing single OAM mode. In addition, it should be noted that, in this paper, we discuss the main factors affecting the recognition accuracy

of applying a six-layer of CNN to OAM communication system, but the detailed structure of CNN itself, such as network internal parameters including the number of network layer, the learning rate, the selection of activation function, the size and number of convolution filter, etc., will also have some impact on the recognition results. These have been extensively discussed in the literature on computer engineering in the past [37]–[41], so we will not discuss them in detail here. All in all, compared with the traditional method, CNN has high detection accuracy of OAM beams, and does not require complicated equipment support and beam calibration at the receiver. Thus, this efficient detection method can be applied to the realization of high-capacity OAM-based optical communication technology in the future.

4. Conclusion

In this paper we focus on the use of CNN in OAM based optical communication systems to recognize different OAM modes under turbulence environment. We expound the corresponding principles of vortex beam, turbulence model and CNN. Then, a specially designed six-layer CNN network in CPU station is established, from which we have investigated the effects of different turbulence-intensity levels, transmission distances and mode spacings on OAM recognition accuracy. Finally, we analyze the recognizing performances of different cases with single OAM mode and multiplexed OAM mode, and briefly discuss the factors that affect the measurement results. The numerical simulation shows that based on our designed CNN structure, through the feature extraction of the received beams' intensity distributions, the recognition of different OAM modes in turbulent environment has achieved efficiently, in which the recognizing accuracy about 96.25% of coaxial multiplexed OAM modes is higher than that of single OAM mode even under long transmission distance with strong turbulence. Due to the high efficiency of the method and no need of redundant optical devices, this efficient detection method might be helpful to future implementation of high-capacity OAM-based optical communication technology.

It should be mentioned that the designed CNN in this paper aims to make a trade-off between the computational complexity of the system and the efficiency of recognition. It still belongs to the category of classic network and there are other network models that can be chosen, such as ZFNet [38], GoogleNet [39], VGGNet [40] and ResNet [41], which are used in different cases in the field of computer engineering and can be further investigated in our next work. We are also planning on studying the applicability of the CNN-based method to LG beams with different radial indices or other OAM carrying beams including Bessel-Gauss (BG) beams. In addition, the combination of different adaptive demodulation algorithms and CNN may be one of the directions for further enhancing the recognition efficiency, which will be considered in our future investigations.

References

[1] A. Banerjee *et al.*, "Wavelength-division-multiplexed passive optical network (WDM-PON) technologies for broadband access: A review," *J. Opt. Netw.*, vol. 4, no. 11, pp. 737–758, 2005.
[2] A. Gnauck, P. Winzer, S. Chandrasekhar, X. Liu, B. Zhu, and D. Peckham, "Spectrally efficient long-haul WDM transmission using 224-Gb/s polarization-multiplexed 16-QAM," *J. Lightw. Technol.*, vol. 29, no. 4, pp. 373–377, Feb. 2011.
[3] A. Gupta, K. S. Bhatia, and H. Kaur, "Comparison of QAM and DP-QPSK in a coherent optical communication system," *Optik-Int. J. Light Electron. Opt.*, vol. 125, no. 20, pp. 5940–5942, 2014.
[4] D. Richardson, J. Fini, and L. E. Nelson, "Space-division multiplexing in optical fibres," *Nature Photon.*, vol. 7, no. 5, pp. 354–362, 2013.
[5] L. Allen, M. W. Beijersbergen, R. Spreeuw, and J. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A*, vol. 45, no. 11, 1992, Art. no. 8185.
[6] J. Wang *et al.*, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nature Photon.*, vol. 6, no. 7, pp. 488–496, 2012.
[7] G. Gibson *et al.*, "Free-space information transfer using light beams carrying orbital angular momentum," *Opt. Exp.*, vol. 12, no. 22, pp. 5448–5456, 2004.
[8] Z. Guo, Z. Wang, M. I. Dedo, and K. Guo, "The orbital angular momentum encoding system with radial indices of Laguerre–Gaussian beam," *IEEE Photon. J.*, vol. 10, no. 5, Oct. 2018, Art. no. 7906511.
[9] C. Kai *et al.*, "The performances of different OAM encoding systems," *Opt. Commun.*, vol. 430, pp. 151–157, 2019.

[10] A. E. Willner *et al.*, "Optical communications using orbital angular momentum beams," *Adv. Opt. Photon.*, vol. 7, no. 1, pp. 66–106, 2015.
[11] T. Lei *et al.*, "Massive individual orbital angular momentum channels for multiplexing enabled by dammann gratings," *Light, Sci. Appl.*, vol. 4, no. 3, 2015, Art. no. e257.
[12] Z. Wang, N. Zhang, and X.-C. Yuan, "High-volume optical vortex multiplexing and de-multiplexing for free-space optical communication," *Opt. Exp.*, vol. 19, no. 2, pp. 482–492, 2011.
[13] N. Zhang, X. Yuan, and R. Burge, "Extending the detection range of optical vortices by dammann vortex gratings," *Opt. Lett.*, vol. 35, no. 20, pp. 3495–3497, 2010.
[14] C. Kai, P. Huang, F. Shen, H. Zhou, and Z. Guo, "Orbital angular momentum shift keying based optical communication system," *IEEE Photon. J.*, vol. 9, no. 2, 2017, Art. no. 7902510.
[15] G. C. Berkhout, M. P. Lavery, J. Courtial, M. W. Beijersbergen, and M. J. Padgett, "Efficient sorting of orbital angular momentum states of light," *Phys. Rev. Lett.*, vol. 105, no. 15, 2010, Art. no. 153601.
[16] J. A. Anguita, M. A. Neifeld, and B. V. Vasic, "Turbulence-induced channel crosstalk in an orbital angular momentum-multiplexed free-space optical link," *Appl. Opt.*, vol. 47, no. 13, pp. 2414–2429, 2008.
[17] S. Zhao, J. Leach, L. Gong, J. Ding, and B. Zheng, "Aberration corrections for free-space optical communications in atmosphere turbulence using orbital angular momentum states," *Opt. Exp.*, vol. 20, no. 1, pp. 452–461, 2012.
[18] Y. Ren *et al.*, "Correction of phase distortion of an OAM mode using GS algorithm based phase retrieval," in *Proc. Conf. Lasers Electro-Opt.*, 2012, pp. CF3I–4.
[19] Y. Ren *et al.*, "Adaptive optics compensation of multiple orbital angular momentum beams propagating through emulated atmospheric turbulence," *Opt. Lett.*, vol. 39, no. 10, pp. 2845–2848, 2014.
[20] D. Wang *et al.*, "Combatting nonlinear phase noise in coherent optical systems with an optimized decision processor based on machine learning," *Opt. Commun.*, vol. 369, pp. 199–208, 2016.
[21] D. Wang *et al.*, "Nonlinear decision boundary created by a machine learning-based classifier to mitigate nonlinear phase noise," in *Proc. Eur. Conf. Opt. Commun.*, 2015, pp. 1–3.
[22] D. Wang *et al.*, "System impairment compensation in coherent optical communications by using a bio-inspired detector based on artificial neural network and genetic algorithm," *Opt. Commun.*, vol. 399, pp. 1–12, 2017.
[23] D. Wang *et al.*, "Intelligent constellation diagram analyzer using convolutional neural network-based deep learning," *Opt. Exp.*, vol. 25, no. 15, pp. 17 150–17 166, 2017.
[24] M. Krenn *et al.*, "Communication with spatially modulated light through turbulent air across vienna," *J. Phys.*, vol. 16, no. 11, 2014, Art. no. 113028.
[25] M. Krenn *et al.*, "Twisted light transmission over 143 km," *Proc. Nat. Acad. Sci.*, vol. 113, no. 48, pp. 13 648–13 653, 2016.
[26] E. Knutson, S. Lohani, O. Danaci, S. D. Huver, and R. T. Glasser, "Deep learning as a tool to distinguish between high orbital angular momentum optical modes," in *Proc. Opt. Photon. Inf. Process. X*, 2016, vol. 9970, Art. no. 997013.
[27] J. Li, M. Zhang, and D. Wang, "Adaptive demodulator using machine learning for orbital angular momentum shift keying," *IEEE Photon. Technol. Lett.*, vol. 25, pp. 1455–1458, Sep. 2017.
[28] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
[29] J. Li, M. Zhang, D. Wang, S. Wu, and Y. Zhan, "Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication," *Opt. Exp.*, vol. 26, no. 8, pp. 10 494–10 508, 2018.
[30] T. Doster and A. T. Watnik, "Machine learning approach to OAM beam demultiplexing via convolutional neural networks," *Appl. Opt.*, vol. 56, no. 12, pp. 3386–3396, 2017.
[31] S. Lohani and R. T. Glasser, "Turbulence correction with artificial neural networks," *Opt. Lett.*, vol. 43, no. 11, pp. 2611–2614, 2018.
[32] S. Lohani, E. M. Knutson, M. ODonnell, S. D. Huver, and R. T. Glasser, "On the use of deep neural networks in optical communications," *Appl. Opt.*, vol. 57, no. 15, pp. 4180–4190, 2018.
[33] Q. Tian *et al.*, "Turbo-coded 16-ary OAM shift keying FSO communication system combining the CNN-based adaptive demodulator," *Opt. Exp.*, vol. 26, no. 21, pp. 27 849–27 864, 2018.
[34] T. Doster and A. T. Watnik, "Laguerre–gauss and bessel–gauss beams propagation through turbulence: Analysis of channel efficiency," *Appl. Opt.*, vol. 55, no. 36, pp. 10 239–10 246, 2016.
[35] L. C. Andrews and R. L. Phillips, *Laser Beam Propagation Through Random Media*, vol. 152. Bellingham, WA, USA: SPIE, 2005.
[36] R. J. Hill, "Models of the scalar spectrum for turbulent advection," *J. Fluid Mech.*, vol. 88, no. 3, pp. 541–562, 1978.
[37] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
[38] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, Springer, 2014, pp. 818–833.
[39] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.
[40] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent*, 2014, pp. 1–14.
[41] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.