Multidimensional Information Assisted Deep Learning Realizing Flexible Recognition of Vortex Beam Modes

Jiale Zhao, Zijing Zhang, Yiming Li, and Longzhu Cen

***Abstract*—Due to countless orthogonal eigenstates, light beams with orbital angular momentum(OAM) have a large potential information capacity. Recently, deep learning has been extensively applied in recognition of OAM mode. However, previous deep learning methods require a constant distance between laser and receiver. The accuracy will drop quickly if the distance of testing set deviates from the training set. Previous deep learning methods also have difficulty distinguishing OAM modes with positive and negative topological charges. In order to further exploit the huge potential of the countless dimension of state space, we proposed multidimensional information assisted deep learning flexible recognition (MIADLFR) method to make use of both intensity and angular spectrum information for the first time to achieve recognition of OAM modes unlimited by the sign of TC and distance with high accuracy. Also, MIADLFR can reduce the computational complexity significantly and requires much smaller training set.**

***Index Terms*—Orbital angular momentum, atmospheric turbulence, deep learning, optical detection.**

I. Introduction

SINCE Allen *et al.* [1] recognized that vortex beam with phase structure $exp(il\phi)$ carries OAM $l\hbar$ per photon, where l is topological charge(TC) vortex beam has been extensively investigated in optical manipulation [2], imaging [3], optical communication [4]. And because of the fact that can take any integer value, vortex beam has great potential in optical communication [5].

In the past 20 years, there are plenty of progress in techniques for sorting OAM modes. Holograms can be used to transform spiral phase structure, thus it can be a mode specific detector [6]. More efficient sorting can be done with Mach-Zehdner interferometer and a Dove prism in each arm [7]. Theoretically, the efficiency can reach 100%, but sorting N modes requires $N-1$ cascaded interferometers. Berkhout *et al.* [8] demonstrated a very successful method for measuring the orbital angular momentum states of light based on log-polar transformation. With this method we can get angular spectrum of vortex beam using two static optical elements. Recently there are also a number of researches improving its resolution [9]–[13]. In fact, gradually-changing-period gratings [14] and annular gratings [15] etc. can also sort Laguerre-Gaussian (LG) beams, but log-polar transformation seems to be the most intuitive method, as a result of which, we will use log-polar transformation method to extract angular spectrum information in this letter. These methods are useful when light beam is ideal. For vortex beam through atmospheric turbulence, deep learning is usually used to recognize OAM modes.

In recent years, deep learning has been widely applied in computer vision. Convolutional neural network(CNN) as a popular model of deep learning has been applied in recognizing OAM modes [16]–[22]. Zhanwei Liu *et al.* first realized superhigh-resolution recognition of OAM modes with the help of deep learning [16]. Junmin Liu proposed a deep learning based atmospheric turbulence compensation method [19].

However, the images sent into CNN are intensity distributions detected directly. As is known to all, it is the spiral phase structure that gives vortex beam angular momentum. The phase information is lost by the sensor. This could leads to two major problems. Firstly, as LG light propagates, the radius of the beam increases, however, the radius of the beam is a quite important feature of LG beam for CNN since larger radius means larger TC. Thus, when the training set and testing set contains LG light propagate different distances, the accuracy of CNN prediction would decrease compared to the same distance case, as we will show in the following. Secondly, because LG light with the same absolute value of TC however with opposite sign share quite similar intensity profile which is all the information sent into CNN, there is no deep learning method realizing sorting LG light with positive and negative TC efficiently. These two drawbacks can severly block the practical application of light with orbital angular momentum. With log-polar transformation, we are able to exploit the phase information. While many other deep learning methods only made use of intensity directly detected by sensor.

In this letter, we proposed and demonstrated a MIADLFR method to remove constrains of previous deep learning methods

Manuscript received June 20, 2021; revised August 1, 2021; accepted August 12, 2021. Date of publication August 18, 2021; date of current version September 2, 2021. This work was supported in part by National Natural Science Foundation of China under Grants 61701139 and 62075049. *(Corresponding author: Zijing Zhang.)*

Jiale Zhao, Zijing Zhang, and Longzhu Cen are with the Department of Physics, Harbin Institute of Technology, Harbin 150001, China (e-mail: marshmallowzjl@gmail.com; zhangzijing@hit.edu.cn; 1121120119@hit.edu.cn).

Yiming Li is with the College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China (e-mail: yiming@cau.edu.cn).

This article has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

Digital Object Identifier 10.1109/JPHOT.2021.3105500

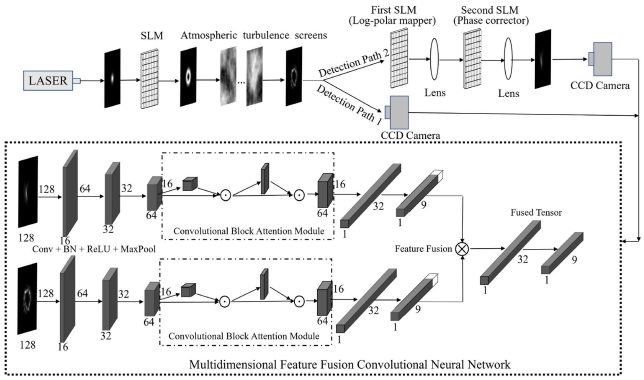

Fig. 1. The system used for MIADLFR method. SLM is spatial light modulator used to transform gaussian beam to LG beam. Atmospheric turbulence screen is for simulating atmospheric turbulence. Log-polar mapper, phase corrector and lens are used to get angular spectrum information. Parameters of MFFCNN is shown in the schematic diagram.

mentioned above. MIADLFR method explores both intensity information and spectrum information at the same time with the help of multidimensional feature fusion convolutional neural network (MFFCNN) proposed in this letter. Log-polar transformation is used to explore phase information. Multidimensional information used together can achieve things impossible for using only intensity information. With multidimensional information, recognition of OAM modes unlimited by distance and sign of TCs can be realized. What's more, MIADLFR can also increase the accuracy of recognition of OAM modes, reduce the size of training set and computational complexity required significantly as we will show in results part. At first, MFFCNN extract features from intensity information and spectrum information. Then fully connected layers process these two dimensions features and gives prediction.

II. Theory

The system used in this letter can be seen in Fig. 1 Firstly, with the help of spatial light modulator (SLM), Gaussian beam generated by laser can be transformed into LG light. Then LG light gets through atmospheric turbulence simulated by atmospheric turbulence screens. There are two paths for detection in system we used. Through the first path, we get the intensity distribution of LG light after atmospheric turbulence. Through the second path, we get the spectrum information with the help of Log-polar transformation method [8]. In the end, angular spectrum information and intensity information are sent into MFFCNN together.

The complex field becomes LG light $U(z,x,y)|_{z=0}$ after SLM, which carries angular momentum. Then it gets through atmospheric turbulence.

For mathematical simplification, treating turbulence as a finite number of discrete atmospheric turbulence screens is a common technique to generate large training sets [19]. A number of numerical models have been proposed to simulate atmospheric turbulence. In this letter we used model developed by Hill [23] and defined by Andrews [24]. The modified von Karman refractive index power spectrum density can be written as:$\phi^2(\kappa) = 0.49r_0^{-5/3}exp(-\kappa^2/\kappa_m^2)/(\kappa^2+\kappa_0^2)^{11/6}$.$r_0$ is the effective coherence diameter given by $r_0 = (0.423k^2C_n^2\Delta z)^{-3/5}$. C_n^2 is the atmospheric refractive index structure constant, representing the turbulence intensity. $\kappa^2 = \kappa_x^2 + \kappa_y^2$, where κ_x and κ_y denote the wavenumber in x and y direction respectively. $\kappa_m = 5.92/l_0$ and $\kappa_0 = 2\pi/L_0$ represent the high-frequency and low-frequency behavior predicted by dimensional analysis. Phase change of atmospheric turbulence screens can be described by $\theta(x,y) = FFT(M\phi(\kappa))$. M is an $N \times N$ dimensional complex random number array with a mean of 0 and a variance of 1. We might as well set $H = exp(ik\Delta z)\cdot exp(-i(\kappa_x^2+\kappa_y^2)\Delta z/(2k))$, which represents Fresnel propagation. Here, Δz is the distance between two atmospheric turbulence screens. Then the beam after going through one atmospheric turbulence screen can be described by:

$$U(z,x,y)|_{z=z_0+\Delta z} = \hat{T}U(z,x,y)|_{z=z_0} \tag{1}$$

$$\hat{T}U(z_0,x,y) = FFT^{-1}[H\cdot FFT[exp(i\theta(x,y))\times U(z_0,x,y)]] \tag{2}$$

Thus, intensity profile after n atmospheric turbulence screens can be represented by:

$$I_1(x,y) = |\hat{T}^n U(z,x,y)|_{z=0}|^2 \tag{3}$$

As a result, the intensity distribution detected in the first path is $I_1(x,y)$.

As shown in Fig. 1 after atmospheric turbulence, first SLM (the Log-polar mapper) can be used to achieve the coordinate transformation $(x,y)\mapsto(u,v)$. Here, $v = a\arctan(y/x)$ and $u = -a\ln(\sqrt{x^2+y^2}/b)$. The phase factor of the Log-polar mapper is given by $\phi_1 = 2\pi a/(\lambda f)[y\arctan(y/x) - x\ln(\sqrt{x^2+y^2}/b)+x]$. Second SLM (the phase corrector) is also needed to eliminate the phase distortion. The phase factor of the phase corrector is given by $\phi_2 = -2\pi ab/(\lambda f)exp(-u/a)cos(v/a)$. Detailed explanation can be seen in ref [8]. For simplicity, the light beam after log-polar transformation is represented as:

$$U_{out} = \hat{L}U_{in} \tag{4}$$

Here, U_{in} is the complex field before log-polar transformation and U_{out} is the complex field after log-polar transformation.Then, angular spectrum profile sent into CNN can be represented by:

$$I_2(x,y) = |\hat{L}\hat{T}^n U(z,x,y)|_{z=0}|^2 \tag{5}$$

As a result, the angular spectrum information detected in the second path is $I_2(x,y)$. As shown in Fig. 1, after log-polar transformation, intensity image and angular spectrum image are sent into two CNN respectively.

Deep learning has merged as an important class of artificial intelligence. Recently, deep learning has shown its great power in image classification [25] and CNN is an important tool in deep learning [26]. CNN is made up with convolutional layers and fully connected layers. Convolutional layers can extract features which are inherently invariant to spatial transformations in images and fully connected layers are able to process information extracted by upstream convolutional layers nonlinearly. After

fully connected layers and appropriate activations, we finally get predicted labels.

The structure of MFFCNN is shown in Fig. 1 What distinguishes MFFCNN from previous CNNs is that the input of MFFCNN contains both intensity and angular spectrum information simultaneously. First, MFFCNN extract features from intensity images and angular spectrum images. In order to ensure the effectiveness of the features, we introduce convolutional block attention module (CBAM) [27] after convolutional layers, which refines features along channel and spatial axes by inferring their attention maps. The obtained tensors $\mathbf{z}_1$, $\mathbf{z}_2$ are considered to exist in different feature spaces and represent individual informative meanings. Next, we use a fusion layer [28] to disentangle dimension specific and cross dimension dynamics by modeling each of them explicitly. The layer is defined as a differentiable outer product between $\mathbf{z}_1$ and $\mathbf{z}_2$:

$$\mathbf{z} = \begin{bmatrix} \mathbf{z}_1 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} \mathbf{z}_2 \\ 1 \end{bmatrix} \tag{6}$$

Here, $\mathbf{z}$ is the fused tensor; $\otimes$ indicates the outer product between tensors. The extra constant dimension with value 1 generates the dimension specific dynamics and thus $\mathbf{z}$ can be viewed as a 2D square of all possible combinations of two tensor spaces.

Finally, the stacked fully connected layers process the fused tensor $\mathbf{z}$ and give final classification results.

Therefore, MFFCNN is used to find a map $f$which can serve as a discriminative boundary among different TCs:

$$\hat{l} = \arg\max f(I_1(x,y), I_2(x,y); W) \tag{7}$$

Here, $\hat{l}$ is the TC predicted by MFFCNN; W represents trainable parameters in MFFCNN; f is the map to find.

In order to find such a map, we need to define loss function to evaluate the difference from prediction and actual TCs. Then through gradient descent we can minimize loss function to make MFFCNN more reliable through mini-batch. The loss function we used can be given by:

$$L = \sum_{i=1}^{M} \sum_{c=0}^{N-1} -y_c^{(i)} \log p_c^{(i)} \tag{8}$$

Here M is the size of testing set; N is the total number of TC in training set; $y_c^{(i)}$ is a binary indicator which takes value 1 if and only if the actual TC of the ith sample of testing set is c. $p_c^{(i)}$ is the probability of the TC of the ith sample to be c which is predicted by MFFCNN.

III. Results

The wave length λ used in this letter is 532 nm. Beam waist w_0 of LG light is 0.03 m. Size of atmospheric turbulence phase screen is 600×600. In simulation $l_0 = 0.0003\text{m}$ and $L_0 = 50$. Before sending images into MFFCNN images are resized to 128×128. In training sets, there are 600 images for each class of OAM modes. In testing sets, there are 200 images for each class of OAM modes. In log-polar transformation, we used following parameters: $a = 0.075/2\pi\text{m}, b = 0.075\text{m}$. Discussion about number of images for each mode in training set is made below.

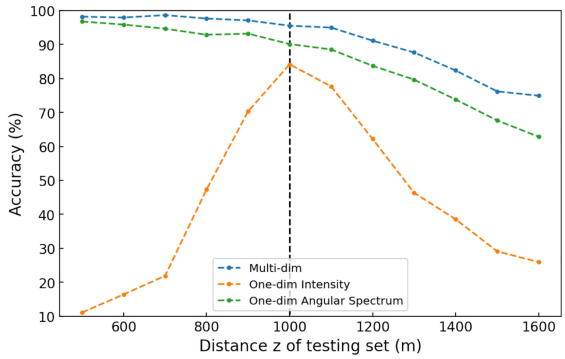

Fig. 2. Accuracy for of multidimensional and one dimensional recognition of OAM modes propagating various distances through atmospheric turbulence of atmospheric refractive index structure constant $C_n^2 = 1 \times 10^{-14}$ with TC ranging from $l = 1$ to $l = 9$. Distance for training set is 1000 m. The vertical line indicates the distance for training set.

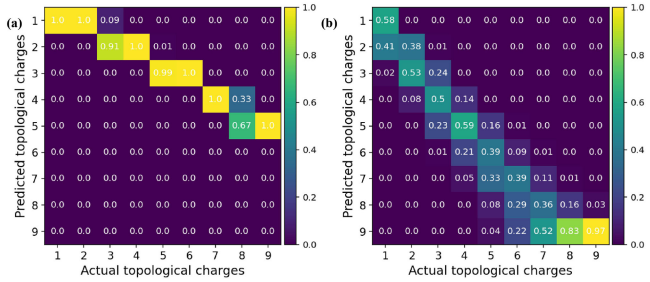

Fig. 3. (a) Intensity one dimensional recognition accuracy when distance for testing set is 200 m. Distance for training set is 1000 m. (b) Intensity one-dim recognition accuracy when distance for testing set is 1600 m.Distance for training set is 1000 m.

A. Recognition of OAM Modes for Arbitrary Distance

LG light with larger TC possess larger beam waist, which makes beam waist a crucial feature extracted by CNN. However, it is known to all that as light beam propagate, the beam waist get larger. As a result, deep learning methods proposed before have difficulty recognizing OAM beams propagating different distances, which is shown in Fig. 2 and Fig. 3 In order to show that unlike previous deep learning methods (Intensity one-dimensional recognition), recognition achieved by MDIADLFR is robust to change of distance, we first use training set generated in a fixed distance(1000 m in our example in Fig. 2) just like previous methods and then change the distance for testing sets.

As is shown in Fig. 2, accuracy for traditional methods only reaches peak when distance of training sets and testing set are the same. The accuracy will drop quickly if distance of testing set is different from training set. This is able to verify our statement above that previous deep learning methods are not suitable for OAM modes recognition for OAM beam propagates distances different from training set. Heat map in Fig. 2 can also further verify our statement. As shown in Fig. 3(a), when the distance of training set is larger than it in testing set, TC predicted by one dimensional CNN is smaller than actual TC, because OAM beam

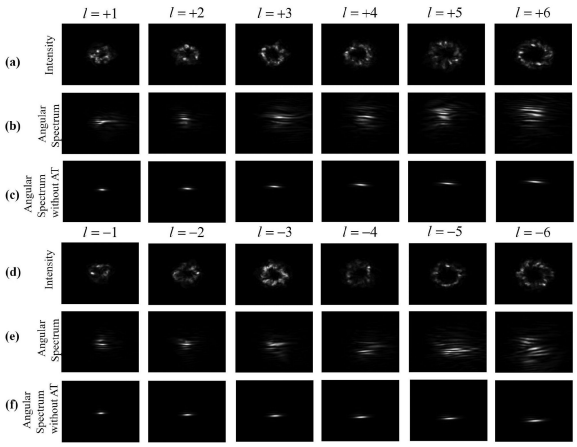

Fig. 4. (a) (d)Intensity images of LG light propagating 1000 m through atmospheric turbulence with C_n^2 value in $1 \times 10^{-14} m^{-2/3}$ carrying various TCs. (b) (e) Angular spectrum images of LG light through atmospheric turbulence with C_n^2 value in $1 \times 10^{-14} m^{-2/3}$ carrying various TCs. (c) (f) Angular spectrum images of LG light without atmospheric turbulence.

with smaller TC have smaller beam waist. On the contrary if distance for training set is smaller than testing set, TC predicted based on one dimensional intensity distribution is larger than actual TC, as we can see in Fig. 3 However, for multidimensional recognition, the accuracy is above 90% for distance less than 1200 m and dips slowly as LG light propagate. Accuracy for angular spectrum one dimensional recognition is lower than multidimensional overall. This seems to because features extracted from these two dimensions compensate each other, as a result of which accuracy for multidimensional recognition is higher than any one dimension.

B. Recognition of OAM Modes Unlimited by Sign of TC

As is known to all that LG lights with the same absolute value of TC but with opposite sign of TC share the same intensity profile. It can be seen in Fig. 4(a) (d) that the intensity profiles of OAM modes with the same absolute value of TC but different sign are quite similar. They share the same radius of beam which is a vital feature CNN extracted. As a result, recognition of OAM light based on intensity distribution alone is not a reliable method. It can be seen in Fig. 4(c)(f) that the transverse position in the fourier plane detected in path 2 is related to the TC. Thus, with angular spectrum we are able to sort OAM beam with the same absolute value of TC but opposite sign of TC efficiently.

Cross talk is shown in the off-diagonal part of the heat map like Fig. 5(a) The cross talk to the same absolute value of TC but with opposite sign of TC is negligible. As shown in Fig. 5(a), MFFCNN is perfectly capable of sorting LG light with positive and negative TC with the same absolute value. The MIADLFR proposed in this paper is capable of recognizing OAM beam with positive and negative TC, which can help further exploit the huge information potential. In addition, as the absolute value of TC grows, the accuracy drops gradually. This is because the cross talk induced by atmospheric turbulence between channels gets severer as the absolute value of TC grows, as shown in Fig. 4(b) (e), which is consistent with previous research [29].

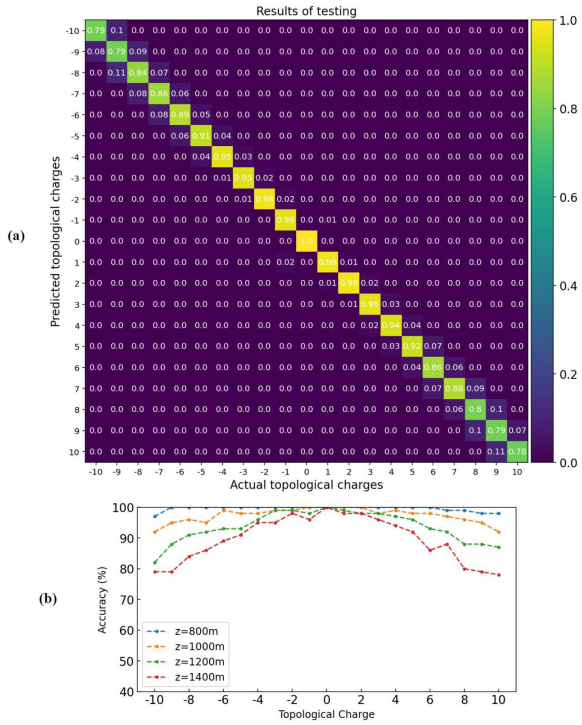

Fig. 5. (a) Recognition accuracy of OAM modes propagating through atmospheric turbulence with C_n^2 value in$1 \times 10^{-14} m^{-2/3}$ and l ranging from -10 to +10. (b) Accuracy for LG beams with various value of TCs at different distances with C_n^2 equals$1 \times 10^{-14} m^{-2/3}$.

C. Accuracy for Various Strength of Atmospheric Turbulence

It is shown in Fig. 6(a) that the accuracy of intensity one dimensional recognition starts to fall when the structure constant of refractive index C_n^2 reaches 1×10^{-15}. For angular spectrum one dimensional recognition, the accuracy of recognition begins to slide when C_n^2 reaches 4×10^{-15}, which shows that angular spectrum alone is also a more accurate way to recognize OAM mode. This is because when the atmospheric turbulence gets stronger, the cross talk between channels tend to be symmetric. For intensity and angular spectrum two dimensional recognition, the accuracy stays over 97.9% even when C_n^2 reaches 7×10^{-15}. Direct detection of intensity distribution can not reflect phase information, which can be extracted by log-polar transformation. Thus information lost in intensity distribution can be compensated by angular spectrum. As a result, with multidimensional information, we can achieve recognition more accurate than any one dimensional recognition.

D. Size of Training Set

As is known to all that larger training set usually result in higher accuracy of course below certain upper limit. Generating training set usually take a lot of time and it takes more time to train a model with larger training set. As a result, we always need to make a trade-off about the size of training set. In most occasions, usually choose points after which the accuracy rises quite slow as the size of training set grows. We might as well refer the size for each class satisfying requirement above as converge point (CP) in this letter.

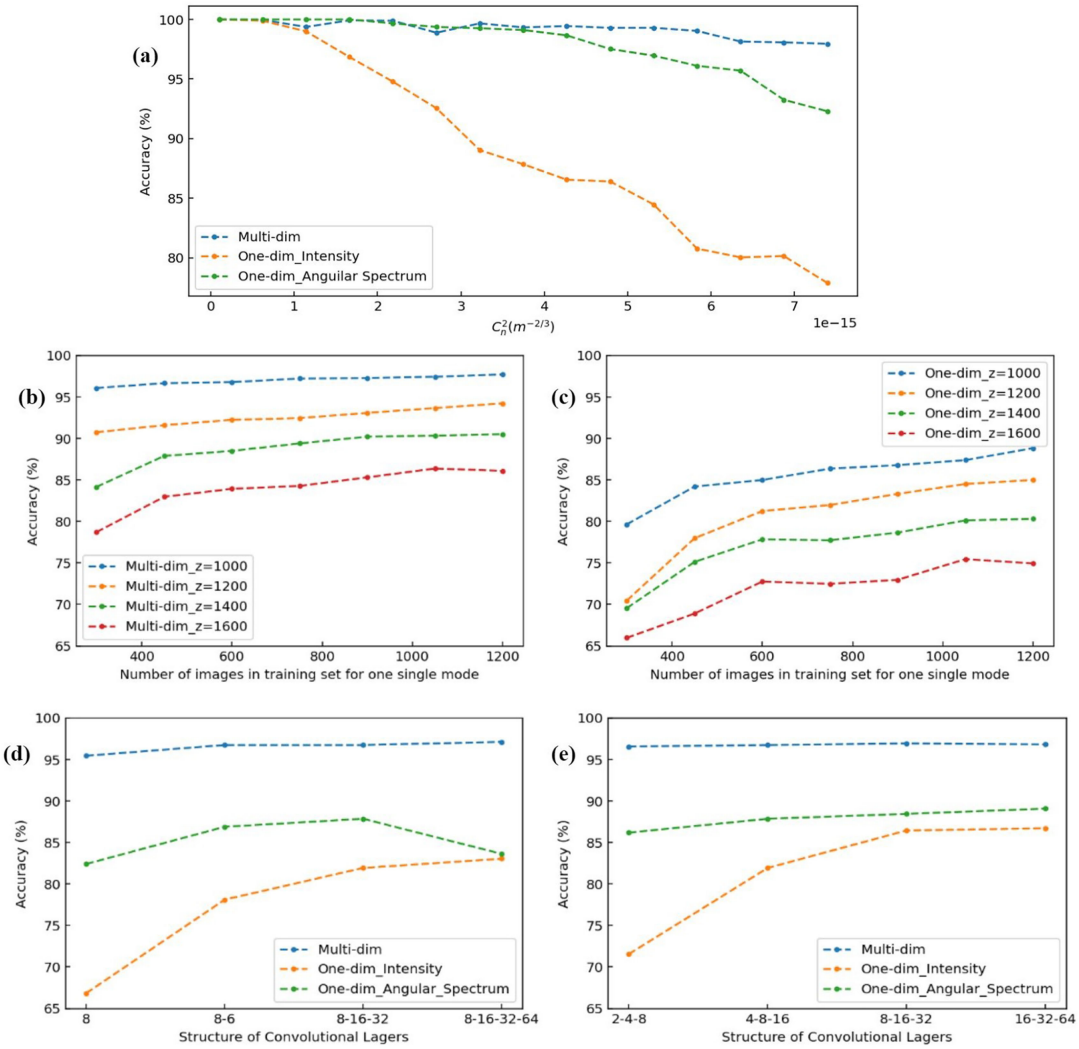

Fig. 6. (a) Accuracy of LG light getting through atmospheric turbulence with various C_n^2 when the distance travelled is 1000 m. (b) Accuracy of MFFCNN recognition for various number of images in training set for one single mode with $C_n^2 = 1 \times 10^{-14}$. (c) Accuracy of recognition for various number of images in training set for one single mode with $C_n^2 = 1 \times 10^{-14}$ using previous methods. (d) Behavior of MFFCNN and one dimensional CNNs with different number of convolutional layers at $z = 1000m$ and $C_n^2 = 1 \times 10^{-14}$. x coordinates represent the structure of convolutional layers. For example, '16-32' means that there are two convolutional layers with 16 and 32 channels respectively. (e) Behavior of MFFCNN and one dimensional CNNs with different number of channels at $z = 1000m$ and $C_n^2 = 1 \times 10^{-14}$.

It can be seen in Fig. 6(b) (c) that the longer LG light propagated that is to say the severer LG light is affected, the larger training set is needed. When LG light propagates 1000 m and 1200 m, the CP is 300 for MFFCNN. Meanwhile, the CPs are 600 and 750 for $z = 1400$ and $z = 1600$ respectively. This result is reasonable, because we usually need larger training set to find best map for classification when the input gets more noise. As for previous methods, CP points are 600 for various distances as shown in Fig. 6(c) In a nut shell, the size of training set required by the MFFCNN is only half that for previous method which can make better use of the potential advantage of large OAM state space.

E. Computation Complexity

From Fig. 6(d) (e), we can see that MFFCNN does not need a complex CNN structure to reach a very high accuracy. In Fig. 6(d), we can also see that MFFCNN can get 95% accuracy even with only one convolutional layer, meanwhile accuracy for one dimensional CNN is only 65%. As we can see one dimensional CNN needs three layers to reach a point after which more complex CNN structure would not raise the accuracy significantly Floating-point operations(FLOPs) is often used to measure the complexity of neural network. The smaller FLOPs the less computational complexity is. The FLOPs needed for one dimensional intensity recognition to reach CP point is 13.9MFLOPs. Meanwhile, that of MFFCNN is only 6.17MFLOP. [30]. We believe this is because angular spectrum dimension information compensates features that need complex CNN to extract from intensity dimension. However, features that cannot be extracted from angular spectrum dimension do not need complex CNN to extract from intensity dimension to extract. As shown in Fig. 6(e), intensity one dimensional recognition needs larger number of channels to reach a stable point, which also means that intensity one dimensional recognition is computationally expensive. What's more, with the help of

intensity dimension information, accuracy for MFFCNN is 10% higher than angular spectrum one dimensional recognition the whole time. To sum up, MFFCNN is much less computationally demanding.

IV. Conclusion

In this letter, we proposed MIADLFR method realize flexible recognition of OAM modes. MIADLFR method can extract features from both intensity and angular spectrum thus it can realize things previous deep learning methods fail to achieve, and raise accuracy of prediction remarkably. Such method can recognize OAM modes propagating different distances and sort OAM modes with positive and negative TCs. When atmospheric turbulence refractive index reaches 7×10^{-5} accuracies are 80.1%, 93.2%, 97.9% for intensity one dimensional, angular spectrum, multidimensional recognition respectively. MFFCNN only need 300 images for each mode in training set to reach a relatively stable point for accuracy. Meanwhile, previous methods need at least 600 images. Smaller training set needed would reduce workload for both generating training set and training remarkably, thus make it easier to exploit the potential advantage of large OAM state space. What's more, MFFCNN needs much simpler CNN structure than previous methods, which makes it much faster to train and recognize OAM modes and thus more practical. With only one convolutional layer, accuracy of MFFCNN can reach 95%. At the same time, accuracy for previous one dimensional intensity recognition is only 65%. This method may have nice cross-fertilization with previous atmospheric turbulence compensation method to further raise the accuracy [19].

MIADLFR method proposed in this letter can not only be used in OAM modes recognition, it can be applied in other problems especially for those require multiple inputs. For some problems, collection of training set is technically or computationally expensive, in which case MFFCNN can also come to use.

V. Disclosures

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in the paper entitled "Multidimensional Information Assisted Deep Learning Realizing Flexible Recognition of Vortex Beam Modes".

References

[1] L. Allen, M. W. Beijersbergen, R. J. C. Spreeuw, and J. P. Woerdman, "Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes," *Phys. Rev. A*, vol. 45, no. 11, pp. 8185–8189, 1992.

[2] L. Paterson, M. P. MacDonald, J. Arlt, W. Sibbett, P. E. Bryant, and K. Dholakia, "Controlled rotation of optically trapped microscopic particles," *Science*, vol. 292, no. 5518, pp. 912–914, 2001.

[3] B. Jack *et al.*, "Holographic ghost imaging and the violation of a Bell inequality," *Phys. Rev. Lett.*, vol. 103, no. 8, 2009, Art. no. 083602.

[4] C. Paterson, "Atmospheric turbulence and orbital angular momentum of single photons for optical communication," *Phys. Rev. Lett.*, vol. 94, no. 15, 2005, Art. no. 153901.

[5] J. Wang *et al.*, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nature Photon.*, vol. 6, no. 7, pp. 488–496, 2012.

[6] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of orbital angular momentum states of photons," *Nature*, vol. 412, no. 6844, pp. 313–316, 2001.

[7] J. Leach, J. Courtial, K. Skeldon, S. M. Barnett, S. Franke-Arnold, and M. J. Padgett, "Interferometric methods to measure orbital and spin, or the total angular momentum of a single photon," *Phys. Rev. Lett.*, vol. 92, no. 1, 2004, Art. no. 013601.

[8] C. G. Gregorius *et al.*, "Efficient sorting of orbital angular momentum states of light," *Phys. Rev. Lett.*, vol. 105, no. 15, 2010, Art. no. 153601.

[9] M. Mirhosseini, M. Malik, Z. Shi, and R. W. Boyd, "Efficient separation of the orbital angular momentum eigenstates of light," *Nature Commun.*, vol. 4, no. 1, p. 2781, 2013.

[10] C. Wan, J. Chen, and Q. Zhan, "Compact and high-resolution optical orbital angular momentum sorter," *APL Photon.*, vol. 2, no. 3, 2017, Art. no. 031302.

[11] S. Zheng *et al.*, "Improve polarization topological order sorting with the diffractive splitting method," *Opt. Lett.*, vol. 44, no. 4, pp. 795–798, 2019.

[12] Y. Wen *et al.*, "Compact and high-performance vortex mode sorter for multi-dimensional multiplexed fiber communication systems," *Optica*, vol. 7, no. 3, pp. 254–262, 2020.

[13] S. Lightman, G.R. HurvitzGvishi, and A. Arie, "Miniature wide-spectrum mode sorter for vortex beams produced by 3D laser printing," *Optica*, vol. 4, no. 6, pp. 605–610, 2017.

[14] K. Dai, C. Gao, L. Zhong, Q. Na, and Q. Wang, "Measuring OAM states of light beams with gradually-changing-period gratings," *Opt. Lett.*, vol. 40, no. 4, pp. 562–565, 2015.

[15] S. Zheng and J. Wang, "Measuring orbital angular momentum (OAM) states of vortex beams with annular gratings," *Sci. Rep.*, vol. 7, no. 1, 2017, Art. no. 40781.

[16] Z. Liu, S. Yan, H. Liu, and X. Chen, "Superhigh-resolution recognition of optical vortex modes assisted by a deep-learning method," *Phys. Rev. Lett.*, vol. 123, no. 18, 2019, Art. no. 183902.

[17] Z. Wang *et al.*, "Efficient recognition of the propagated orbital angular momentum modes in turbulences with the convolutional neural network," *IEEE Photon. J.*, vol. 11, no. 3, Jun. 2019, Art. no. 7903614.

[18] T. Doster and Abbie T. Watnik, "Machine learning approach to OAM beam demultiplexing via convolutional neural networks," *Appl. Opt.*, vol. 56, no. 12, pp. 3386–3396, 2017.

[19] J. Liu *et al.*, "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," *Opt. Exp.*, vol. 27, no. 12, pp. 16671–16688, 2019.

[20] 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. 10494–10508, 2018.

[21] Y. Zhai, S. Fu, J. Zhang, X. Liu, H. Zhou, and C. Gao, "Turbulence aberration correction for vector vortex beams using deep neural networks on experimental data," *Opt. Exp.*, vol. 28, no. 5, pp. 7515–7527, 2020.

[22] Z. Li, J. Su, and X. Zhao, "Atmospheric turbulence compensation with sensorless AO in OAM-FSO combining the deep learning-based demodulator," *Opt. Commun.*, vol. 460, 2020, Art. no. 125111.

[23] R. J. Hill, "Models of the scalar spectrum for turbulent advection," *J. Fluid Mechanics*, vol. 88, no. 3, pp. 541–562, 1978.

[24] L. Andrews and R. Phillips, *Laser Beam Propagation Through Random Media*. Bellingham, WA, USA: SPIE, 2005.

[25] 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.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[27] S. Woo, J. Park, Joon-Young Lee, and In So Kweon, "CBAM: Convolutional Block Attention Module," in *Proc. Eur. Conf. Comput. Vis.*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., Cham, Switzerland: Springer, 2018, vol. 11211, pp. 3–19.

[28] A. Zadeh, M. Chen, S. Poria, E. Cambria, and L.-P. Morency, "Tensor fusion network for multimodal sentiment analysis," in *Proc. Conf. Empirical Methods*, 2017.

[29] Jaime A. Anguita, Mark A. Neifeld, and Bane 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.

[30] P. Molchanov, S. Tyree, T. Karras, T. Aila, and J. Kautz, "Pruning convolutional neural networks for resource efficient inference," in *Proc. 5th Int. Conf. Learn. Representations*, Toulon, France, 2017. Online. [Available]: https://openreview.net/forum?id=SJGCiw5gl.