scientific reports

OPEN

Deep-learning-based high-resolution recognition of fractional-spatial-mode-encoded data for free-space optical communications
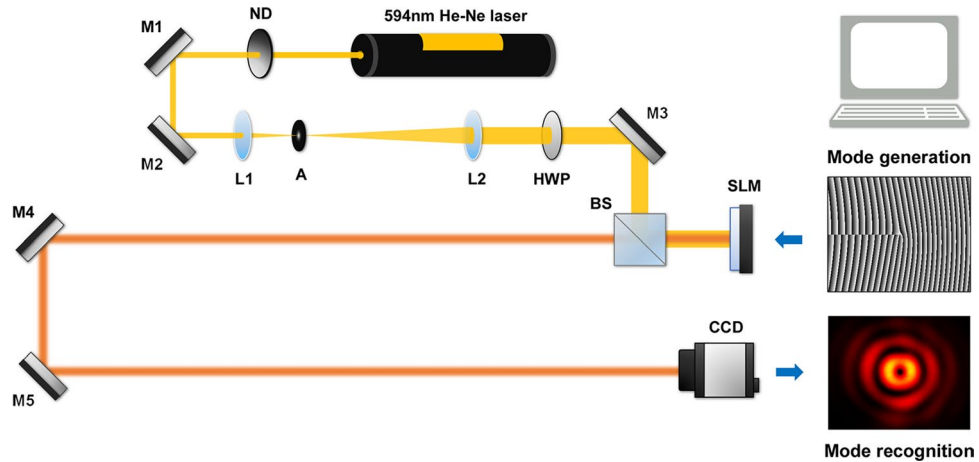
Youngbin Na & Do-Kyeong Ko✉

Structured light with spatial degrees of freedom (DoF) is considered a potential solution to address the unprecedented demand for data traffic, but there is a limit to effectively improving the communication capacity by its integer quantization. We propose a data transmission system using fractional mode encoding and deep-learning decoding. Spatial modes of Bessel-Gaussian beams separated by fractional intervals are employed to represent 8-bit symbols. Data encoded by switching phase holograms is efficiently decoded by a deep-learning classifier that only requires the intensity profile of transmitted modes. Our results show that the trained model can simultaneously recognize two independent DoF without any mode sorter and precisely detect small differences between fractional modes. Moreover, the proposed scheme successfully achieves image transmission despite its densely packed mode space. This research will present a new approach to realizing higher data rates for advanced optical communication systems.

Since the beginning of innovations in information technology, such as the Internet of Things, big data, cloud computing, and artificial intelligence, demand for high-capacity communication systems has been explosively growing. Despite significant improvements in the capacity by the use of wavelength- and polarization-division multiplexing techniques and the application of multilevel modulation formats[1–4], exponentially growing data traffic is facing bandwidth crunch. Light orbital angular momentum (OAM), which is one of the spatial degrees of freedom (DoF), has been proposed as a potential solution to overcome the limitation[5,6]. Light beams with a helical wavefront originating from an azimuthal phase $\exp(im\varphi)$ have an OAM of $m\hbar$ per photon, where m is a topological charge (TC), φ is the transverse azimuth angle, and $\hbar$ is the reduced Planck constant[7,8]. Due to the theoretically unlimited integer values of m, OAM modes potentially provide an infinite-dimensional space, either as data symbols representing M-ary numbers[9–13] or as independent information carriers for multiplexing[14–19]. Indeed, many results have shown that it is possible to achieve data capacity of Tbit/s or up to Pbit/s level through OAM mode division multiplexing in conjunction with other multiplexing techniques, over both free space and fibers[14–16]. However, contrary to the expectation, OAM does not increase the total amount of information and does not outperform both conventional line-of-sight multi-input multi-output transmission and spatial mode multiplexing with a complete basis[10,20,21]. In fact, OAM is only a subset of the transverse laser spatial modes, and the number of available spatial modes is restricted by the space-bandwidth product of a given optical system[10,20,21]. Hence, effective solution is to use full spatial DoF, e.g., both radial and azimuthal modes of Laguerre-Gaussian beams[10,11], and one could consider separating the modes into a fractional interval to increase the addressable mode number in the limited system.

Unlike physical quantities defined on a continuous parameter space, such as linear momentum, OAM eigenmodes form a discrete set of modes where m takes only integer values due to the periodic nature of the angle variable[22]. However, there also exist modes with arbitrary fractional OAM, and it can be described by a coherent superposition of integer OAM modes[22–24]. In other words, despite the OAM range limited by the system space-bandwidth product, one can employ more OAM modes by reducing the OAM interval to a fraction. Various methods have been proposed for measuring and recognizing OAM modes separated by a fractional interval.

Department of Physics and Photon Science, Gwangju Institute of Science and Technology, Gwangju 61005, Republic of Korea. ✉email: dkko@gist.ac.kr

Figure 1. Top view of experimental setup: He–Ne laser of 594 nm wavelength; *ND* neutral density filter; *M* mirror; *L* lens; *A* aperture; *HWP* half wave plate; *BS* beam splitter cube; *CCD* charge coupled device. Propagation distance from the SLM to the CCD camera is 2.8 m. Each inset shows a hologram for mode generation and a captured intensity image for mode recognition.

Some representative ways are based on specifically designed phase elements, which range from a conventional fork hologram[25,26] to an annular grating[27], multifocal arrays[28], and mode sorters utilizing optical coordinate transformation[29,30]. These methods can obtain quantitative values of fractional OAM by measuring either intensity or focal spot displacement, but there are drawbacks to be improved, such as a low resolution, precision, and a need for additional optical elements. Furthermore, these methods are severely affected by optical alignment. Lateral displacement of a demodulation element relative to the optical axis leads to the power leakage into the neighboring modes, altering the OAM spectrum[31]. These issues make it difficult to develop a reliable optical system for fractional OAM beams.

Recently, a convolutional neural network (CNN), the most preferred solution for image classification[32,33], has drawn attention as an efficient tool for recognizing OAM modes and correcting phase distortion caused by atmospheric turbulence[34–39]. End-to-end recognition of the deep-learning allows the classification process to be performed with only the intensity profile of the target modes, which is the most distinctive feature compared to the existing methods that require additional components to extract phase information[35]. Translation invariance of CNN provides stable recognition performance regardless of lateral displacement of a detector to the optical axis, and even it is possible to impart several transform invariances, such as rotation and scaling, through preparation and augmentation of proper data samples[35,39–41]. Moreover, a recent study has experimentally demonstrated that fractional OAM modes can be precisely recognized with a resolution of 0.01[41].
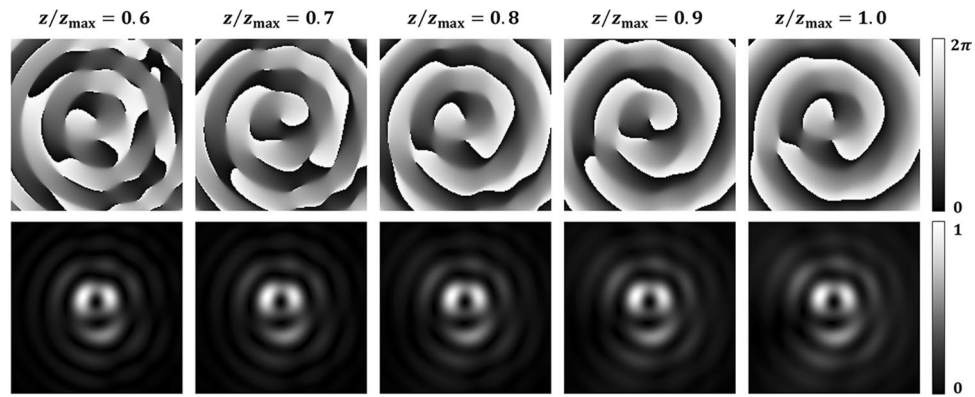
In this paper, we propose fractional modulation of laser spatial modes effectively to increase the available mode number in limited optical systems. In particular, a CNN decoder is applied to achieve high-resolution recognition of fractional modes. The generation and modulation of spatial modes carrying a data symbol are implemented by switching phase holograms displayed on a spatial light modulator (SLM). The transmitted laser modes are captured by a CCD camera and directly recognized in real time by the trained CNN. As a proof of concept, gray values are encoded in 256 spatial modes given by a combination of radial modes and OAM modes and then transmitted in the proposed optical link. Despite the extremely small differences between the fractional modes, our CNN decoder successfully recovers the transmitted data without any optical mode sorter like a fork hologram. Using the proposed deep-learning-based technique, we were able to achieve recognition accuracy higher than 99% for the test dataset and transmit images with correlations higher than 0.99 and error rates lower than 0.10%.

Experimental setup and methods

Experimental setup. Figure 1 schematically depicts the experimental setup to generate and recognize spatial modes of BG beams. A light beam emerging from a He–Ne laser of 594 nm wavelength is expanded and collimated by lenses L1 ($f = 50$ mm) and L2 ($f = 150$ mm). Then the laser beam of 2.1 mm radius is horizontally polarized through a half-wave plate and incident on an SLM. The SLM (PLUTO-VIS, Holoeye) is a reflective phase-only device with 1920×1080 pixels (8 μm pixel pitch). A spatial mode of BG beams is generated by the corresponding phase hologram displayed on the SLM. After 2.8 m propagation from the SLM, the generated BG beam is captured by a CCD camera of 768×576 pixels (8.3 μm pixel pitch). The captured intensity image is downsampled from 400×400 to 100×100 pixels. The down sampled image is fed to a neural network as an input, and the transmitted laser modes are predicted by the network in real time.

Fractional modes of BG beams. Phase holograms for generating BG beams are given by[42]

$$\Phi_{k_r,m}(r,\varphi) = \exp\left[im\varphi + ik_r r\right], \tag{1}$$

Figure 2. Numerically calculated phase and intensity profiles of a fractional BG beam at different positions. Propagation distance normalized to the non-diffracting distance is displayed on the top of the figure, and used parameters are as follows: $w = 2.1$ mm, $\lambda = 594$ nm, $k_r = 7.10$, and $l = 1.54$.

where m is the azimuthal index that determines the order of the BG beam and a TC associated with light OAM. k_r is the radial wavenumber that determines the spacing of the intensity rings and the non-diffracting distance given by[42]

$$z_{\max} = w\frac{k}{k_r},\qquad(2)$$

where w is the radius of an input beam, and $k = 2\pi/\lambda$ is the propagation constant. Various BG modes can be generated by combining mode indices m and k_r in Eq. (1). Unlike the radial modes of which the mode index is a continuous parameter, a fractional OAM mode is given by a coherent superposition of all possible integer modes, not a single eigenmode, due to the discrete parameter m[22,24]. Spiral phase with a fractional index l is expressed as[23]

$$\exp(il\varphi) = \sum_m c_m \exp(im\varphi) = \frac{\exp(il\pi)\sin(l\pi)}{\pi}\sum_m \frac{\exp[im(\varphi+\phi)]}{l-m},\qquad(3)$$

where the complex coefficient c_m of each integer mode is calculated by the orthonormal condition, and ϕ is a phase shift parameter that determines the orientation of the phase discontinuity and results in the rotation of the intensity profile; see Fig. 4b. Theoretically, the decomposition of fractional OAM modes includes all integer modes. However, the optical realization of the modes is restricted by the size of a physical aperture[24] and above all, most of energy is distributed to adjacent 8 modes, e.g., $l = -2$ to 5 in our case. In fact, the center ring structure of fractional beams, which provides crucial local features for each mode, consists of theses modes. The field distribution produced by the phase hologram of Eq. (1) can be described by the Fresnel diffraction integral[42]
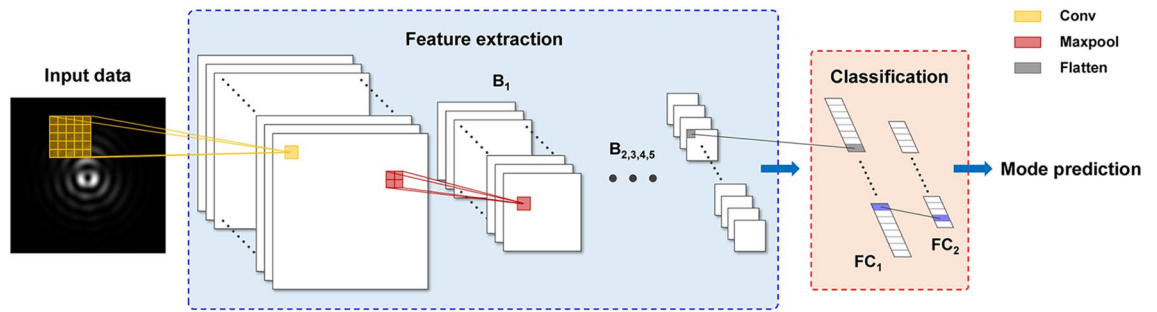
$$
\begin{aligned}
u_{k_r,m}(\rho,\theta,z) &= \frac{1}{i\lambda z}\exp\left[ik\left(z+\frac{\rho^2}{2z}\right)\right]\iint u_g(r,\varphi,0)\Phi_{k_r,m}(r,\varphi) \\
&\times \exp\left(i\frac{kr^2}{2z}\right)\exp\left[-i\frac{k\rho r}{z}\cos(\varphi-\theta)\right]rdrd\varphi \\
&\approx \exp\left[im\left(\theta-\frac{\pi}{2}\right)\right]\int_0^\infty \exp\left[-\frac{r^2}{w^2}-ik_r+i\frac{kr^2}{2z}\right]J_m\left(\frac{k\rho}{z}r\right)rdr,
\end{aligned}\qquad(4)
$$

where z is the propagation distance, and $u_g = \exp(-r^2/w^2)$ is the field distribution of an incident Gaussian beam. Here, we omitted the first term in the first line for the sake of simplicity and used the integral form of the high-order Bessel function[43]. Thus the resulting fractional BG beams with a radial index k_r and an azimuthal index l is given by

$$U_{k_r,l}(\rho,\theta,z) = \sum_m c_m u_{k_r,m}(\rho,\theta,z).\qquad(5)$$

The Gouy phases of different OAM modes give rise to the unstable evolution of light beams emerging from a fractional spiral phase plate[24]. However, after a certain propagation distance, after $0.6z_{\max}$ in our case, the transverse intensity profile of the mode does not change, as shown in Fig. 2. Besides, the fractional BG beam is diffraction-free during propagation over the finite distance, which indicates that the propagation properties of the ordinary BG beams are applied to Bessel beams carrying fractional OAM. Unlike integer OAM beams with a doughnut shape, distinctive structures of the fractional OAM modes provide various local features and allow CNN easily to discriminate differences between adjacent modes.

To demonstrate the feasibility of the densely encoded data transmission via a deep-learning method, we selected a total of 256 modes given by a combination of radial modes from $k_r = 7.10$ to $k_r = 7.80$ with $\Delta k_r = 0.10$ (8 modes) and OAM modes from $l = 1.03$ to $l = 1.96$ with $\Delta l = 0.03$ (32 modes). Meanwhile, a blazed phase grating and a surface correction hologram, made by a combination of Zernike polynomials, are additionally

Figure 3. Structure of the designed CNN: B, block; FC, fully connected layer. A series of blocks which consist of a convolutional layer and a max pooling layer extract features describing the input object. The extracted features are classified through two FC layers, also referred to as dense layers. Finally, a spatial mode of the input is determined by a softmax activation function included in the last FC layer.
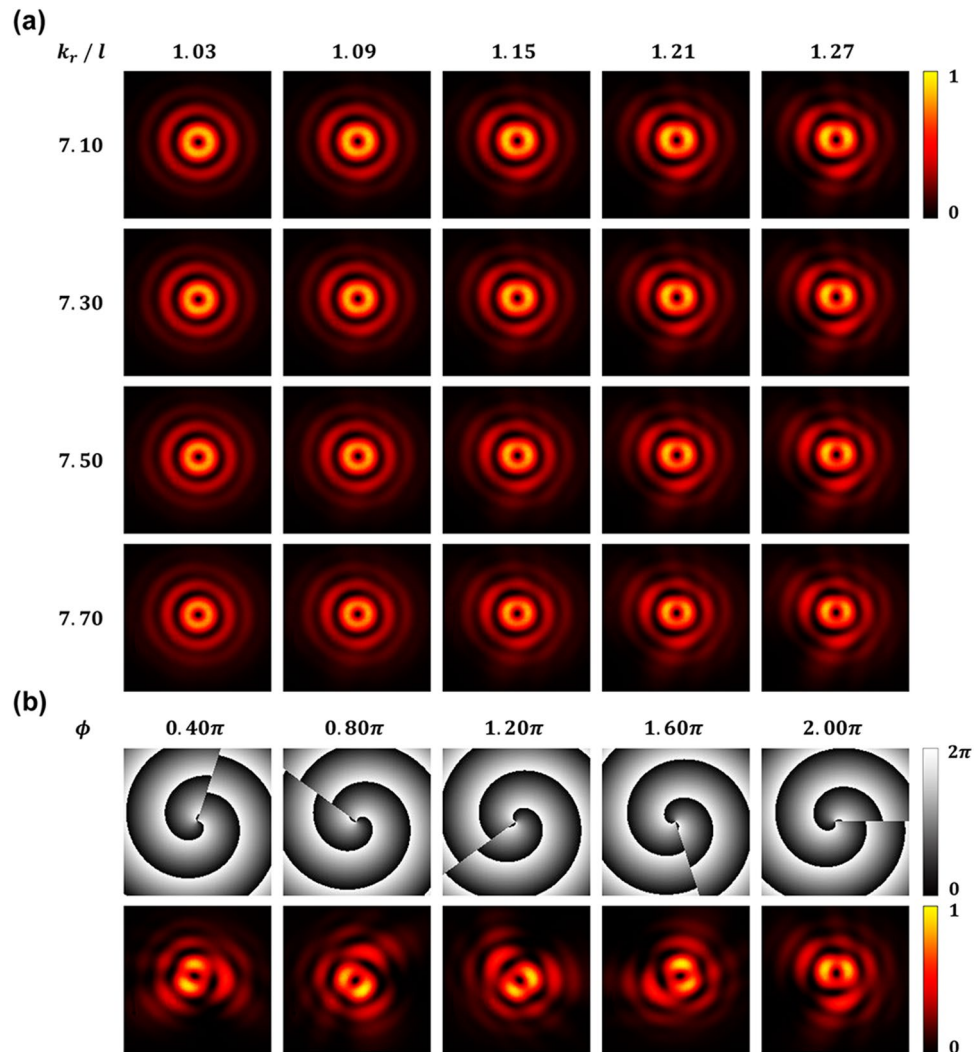
	Layer	Setting	Output shape	# of parameters
B1	Conv2D_1	64, (5, 5), ReLU, padding	(100, 100, 64)	1,664
	Maxpool2D_1	(2, 2)	(50, 50, 64)	0
B2	Conv2D_2	64, (4, 4), ReLU, padding	(50, 50, 64)	65,600
	Maxpool2D_2	(2, 2)	(25, 25, 64)	0
B3	Conv2D_3	128, (3, 3), ReLU, padding	(25, 25, 128)	73,856
	Maxpool2D_3	(2, 2)	(12, 12, 128)	0
B4	Conv2D_4	128, (3, 3), ReLU, padding	(12, 12, 128)	147,584
	Maxpool2D_4	(2, 2)	(6, 6, 256)	0
B5	Conv2D_5	256, (3, 3), ReLU, padding	(6, 6, 256)	295,168
	Maxpool2D_5	(2, 2)	(3, 3, 256)	0
–	Flatten	–	(2304)	0
–	Dropout	5%	(2304)	0
FC1	Dense_1	512, ReLU	(512)	1,180,160
FC2	Dense_2	256, softmax	(256)	131,328

Table 1. Summarized information on the designed CNN.

included in Eq. (1). The former separates the desired first diffraction order, and the latter compensates for the phase deformation caused by the spatial inhomogeneity of the SLM.

Designed network structure and dataset preparation. The designed CNN model comprises two parts for feature extraction and classification, as depicted in Fig. 3. The part of the feature extraction is constructed by 5 blocks, and each block consists of a convolutional layer and a max pooling layer. In this experiment, the number of blocks was determined considering the number of trainable parameters and model performance. The number of parameters in the first fully connected (FC) layer is proportional to the pixel size of input feature maps. Therefore, a model causes inefficiently large number of parameters if the size of feature maps is not scaled down properly. For this reason, 4 and 5 block models were taken into account, and the latter showed the lower validation loss despite similar training times. Multiple convolution kernels in the convolutional layer detect various local features of input images[32]. Then the extracted feature maps are processed by a nonlinear activation function. The purpose of the nonlinear activation function is to introduce non-linearity into the output, which allows the network to learn complex data and provide an accurate prediction[33]. Then the size of the activated feature maps is reduced by half through a max pooling layer, which allows the model to detect the weak variance coming from fractional mode separation[41]. The features extracted from the last block are connected to the classification part comprised of two fully connected (FC) layers. Finally, a softmax function included in the last FC layer yields the probabilities that the transmitted mode belongs to specific laser modes, and then a laser mode with the maximum value is determined to be a received mode; see Eq. (7).

Detailed information on the CNN is summarized in Table 1, including the hyperparameters, output shape, and the number of trainable parameters. Each component in the third column of the table represents the number of convolution kernels, kernel size, activation functions, and zero paddings. For FC layers, each component represents the number of hidden units, also referred to as neurons of the layer, and activation functions. Additionally, He initialization useful for networks with the rectified linear unit (ReLU) as a nonlinear activation function was selected to initialize weight parameters[44]. The dropout layer placed in front of the first FC layer randomly sets 5% of its neurons to 0 for each training epoch to prevent overfitting[45]. It is activated only during the training time, which does not affect the inference.

Figure 4. Experimental generation of fractional BG beams. (**a**) Some examples of the measured intensity profiles and (**b**) phase shift modulation of the fractional BG beam with $k_r = 7.10$ and $l = 1.45$.

To train and test the CNN, 51,200 intensity profiles with 200 images per mode are prepared. These 200 images have different phase shifts from 0.01 to 2.00π, respectively. Of the 200 images, 100 images, corresponding to phase shifts $\phi = 0.02, 0.04, \ldots, 2.00\pi$, are used for training the model, and the others for testing. Additionally, 20 images with different phase shifts from 0.05 to 1.95π per mode are prepared as a validation set. The role of the validation set is to provide an unbiased evaluation of the model during the training process. Meanwhile, the test set is used to assess the prediction accuracy of the final model. All images are resized from 400×400 to 100×100 pixels for computational efficiency and preprocessed for supervised learning.

Experimental results and discussion

Measurement of generated BG beams. Figure 4a displays captured intensity profiles of fractional BG beams generated with different mode indices. Here, pseudo-color, which varies smoothly from black to red and then yellow, is used for visual clarity, but the format of actual images is 8-bit grayscale. Only a few of all 256 spatial modes are presented to show and highlight the tiny variation caused by the fractional mode separation. As the radial wavenumber increases, the spacing between intensity rings is reduced. The intensity of the center ring decreases while the outer ring slightly brightens. Meanwhile, a spiral phase with a fractional TC results in exotic intensity distribution. As the charge increases from $l = 1$, the central ring of high-intensity forms two brighter light petals, and the second ring is transformed into a relatively dark and distorted four-petal structure. The angular position of those intensity peaks depends on the applied phase shift. The lower one of the four petals gradually approaches the central ring. Although not included in Fig. 4a, the petal moves further inward and then merges with the center ring. Finally, the intensity profile becomes a multiple ring structure again at $l = 2$. Using various convolution kernels, the deep-learning model detects invisible variations in local intensity distribution caused by the fractional mode separation. Then it gives the prediction, with high accuracy, based on the decision boundary formed by training data.

Figure 4b depicts the influence of the applied phase shift on the intensity distribution of fractional BG beams. As described above, the intensity profiles of the generated BG beams are rotated counterclockwise by the amount of the phase shift, which means that phase modulation applied to fractional modes is explicitly recognizable. In other words, unlike the canonical integer vortices with the rotational symmetry, broken symmetry of the fractional modes can provide an additional DoF for information coding. It will be a new kind of phase-shift keying technique that exploits phase-shifted fractional modes as M-ary symbols.

Training of neural network and performance test. The CNN training and prediction are performed on the commercial GPU system (GPU: RTX 2060; CPU: i7-9750H) using Keras framework. The parameters of CNN are trained using the Adam optimizer with batch size 50 for 50 epochs. The Adam optimizer is a stochastic gradient descent method, which is based on adaptive estimation of moments[46]. It is computationally efficient, making it suitable for networks with large parameters. The weight parameters for the spatial mode recognition are updated to minimize a loss (objective) function during the training process. The loss function used for this purpose, i.e., for the multiclass classification problem, is the averaged categorical cross-entropy, which is given by[33,41]

$$L = -\frac{1}{N} \sum_{i=1}^{N} \sum_{j=1}^{n} t_j^{(i)} \log \left(S_j^{(i)} \right), \tag{6}$$

where N is the total number of data to be considered, n is the total number of classes, i.e., representing 256 spatial modes, and index j represents each mode. For example, 1-32th classes correspond to modes with $k_r = 7.10$ and $l = 1.03$ to $l = 1.96$, and next 32 classes correspond to modes with $k_r = 7.20$ and $l = 1.03$ to $l = 1.96$. Meanwhile, $t_j^{(i)}$ is the jth element of a target label vector. The label vector is a vector of size n with one on the element corresponding to a specific spatial mode and zeros elsewhere. For example, in our experiment a label vector of the spatial mode with $k_r = 7.10$ and $l = 1.03$ is $\{1, 0, \ldots, 0\}$, and that of the spatial mode with $k_r = 7.10$ and $l = 1.06$ is $\{0, 1, 0, \ldots, 0\}$. $S_j^{(i)}$ is the output of a softmax activation function, which represents the probability that the ith sample belongs to the jth class, and it is given by[33]
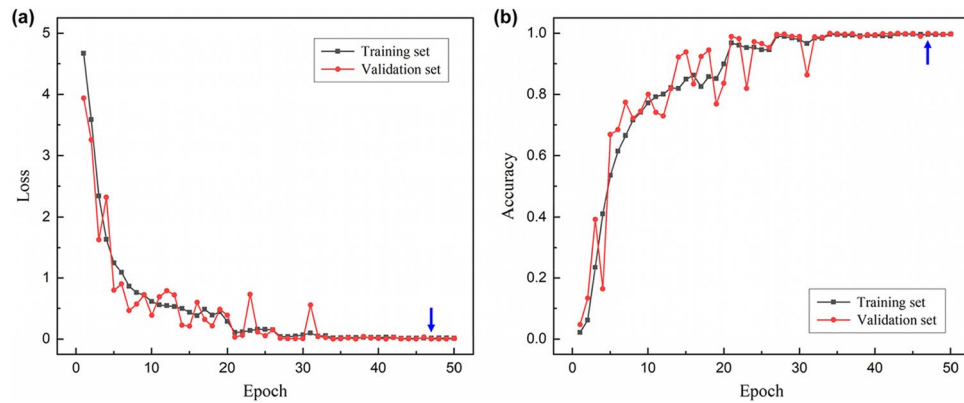
$$S_j = \frac{\exp(y_j)}{\sum_{m=1}^{n} \exp(y_m)}, \tag{7}$$

where y_m is the input of the mth unit, and n is the total number of classes as described above. The initial learning rate is 0.001, which is reduced by half whenever validation loss does not be improved for 5 epochs. After the CNN training, prediction time was measured to be $t_p < 4$ms and it is shorter than the switching time of the SLM. In other words, our CNN model can decode all transmitted data without a delay of signal, which indicates that real-time prediction of laser spatial modes is possible. The calculation speed is affected by the Floating-point Operations Per Second (FLOPS) of the computing system[38]. Here, we employed RTX 2060 of a laptop system, and the FLOPS of a RTX 3080, one of the recent advanced GPU, is about four times higher. It implies that using a more advanced GPU can shorten the prediction time to sub-ms. In other words, the mode detection using the proposed model can be sufficiently applied to devices with fast switching speed, such as a digital micromirror device (DMD). The DMD is a pixelated device of micromechanical mirrors individually switched "on" or "off" states. The device is used as a binary amplitude mask for beam steering and shaping, and a fast refresh rate of up to a few kilohertz or faster is its typical characteristic[47].
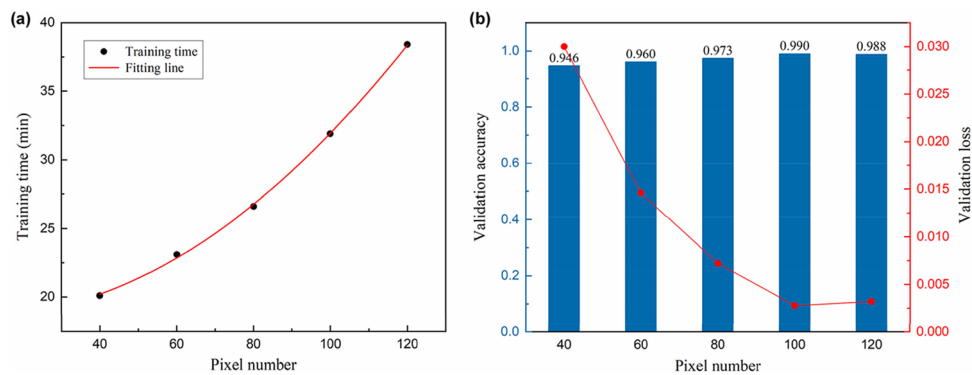
Figure 5 shows the loss and accuracy curves for the training and validation sets as a function of the training epoch. Here the accuracy is defined as the number of correctly classified samples divided by the total number of samples. As shown in Fig. 5a, the model suitably learns the training data and performs well on new, unseen data without overfitting. Note that the validation set never participates in the training process and is used only to evaluate the model every epoch. Minimum validation loss 6.46e−4 was achieved at the 47th epoch, and the validation accuracy was 99.9%. The model and weight parameters that had achieved the minimum validation loss were saved and used for inference.

In addition to the recognition accuracy of the model, computational speed is a critical factor for a communication purpose and real-time prediction. For this reason, we investigated the influence of the input pixel size on the training time and model performance. The image set is resized from 400×400 to 5 different pixel numbers, from 40×40 to 120×120 pixels, using the open-source computer vision library (OpenCV) and fed to the network for training. As can be seen in Fig. 6a, the training time is proportional to the square of the pixel number, which is reasonable when considering that an image is a 2-d matrix. Figure 6b shows the evaluation results obtained with different pixel numbers. A minimum loss of 2.78e-3 and a maximum accuracy of 99.0% were achieved at pixel number 100. Because of a lack of information, the models trained with samples of a resolution lower than 100×100 pixels show poor performance, i.e., underfitting has occurred. On the other hand, the dataset of 120×120 pixels gives rise to the overfitting of the model, which means that the model has learned too many details like noise and may not be able to predict new data in the future. Regularization techniques such as l2 regularization, dropout, and batch normalization can mitigate the overfitting[45,48,49], but using a dataset of 100×100 pixels is more efficient for model training and real-time prediction.

With the trained model described above, an additional test was implemented to demonstrate the classification performance. Figure 7a shows a part of the confusion matrix obtained from the prepared test dataset. The numbers displayed on axes represent spatial modes given by a combination of a radial mode 7.30 and OAM modes from 1.03 to 1.96, which is employed for the sake of simplicity. The diagonal elements indicate that the transmitted mode is correctly classified. Only 25 of the 25,600 images make errors, which are because of the adjacent modes, and the measured test accuracy is 99.9%. The experimental results demonstrate that the trained model can

Figure 5. Training results of the designed CNN. (**a**) Loss and (**b**) accuracy curves for the training set and validation set. Blue arrows indicate the 47th epoch where the validation loss is minimum. Validation loss and validation accuracy were 6.46e−4 and 99.9% at the epoch.
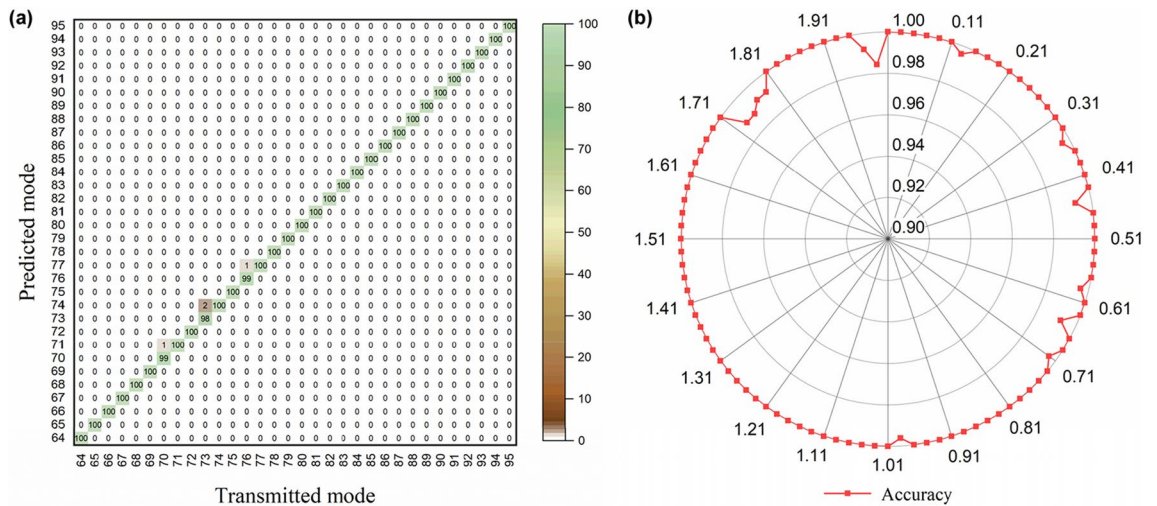

Figure 6. Performance evaluation of the number of input image pixels. (**a**) Time taken to train for 50 epochs and (**b**) validation loss and validation accuracy. All data show mean values of 10 training results per pixel number.

simultaneously identify two independent spatial modes, without any optical mode sorters like a fork hologram. Besides, the model accurately recognizes extremely small differences between fractional modes regardless of the applied phase shift, as can be seen in Fig. 7b. It means that the OAM phase-shift keying proposed in subsection of measurement of generated BG beams could be combined to encode more data without degradation of the recognition accuracy. This is a distinct advantage compared to the optical coordinate transformation method, which shows different output results by the extrinsic OAM component of fractional OAM beams depending on the orientation angle of the phase discontinuity[30].

Image transmission by using spatial mode encoding. As a proof of concept, 8-bit grayscale images are transmitted, pixel by pixel, through free space. 8-bit encoding is performed with 256 spatial modes separated by fractional intervals, which demonstrates the possibility of super dense optical communication assisted by laser spatial modes and deep-learning. As shown in Table 2, 32 OAM modes with the same radial wavenumber are assigned to 32 levels of the 256 gray levels. For example, laser spatial modes ($k_r = 7.10$, $l = 1.03$) and ($k_r = 7.10$, $l = 1.06$) are assigned to 0 and 1, respectively. In other words, 256 different laser spatial modes act as 256-ary symbols.

Image transmission is performed according to the following procedure. First, an image is transformed into a 1-d vector comprised of a series of 256-ary numbers, and the data encoding is done by switching the phase holograms corresponding to each value. At this time, start and stop frames specifying the beginning and end of each data stream are appended every 20 pixels, i.e., 20 symbols, to prevent timing errors between the transmitter and receiver. After free-space propagation, the transmitted laser beam is captured by a CCD camera placed at the receiver, and the trained CNN model decodes the data in real time from the received intensity image.
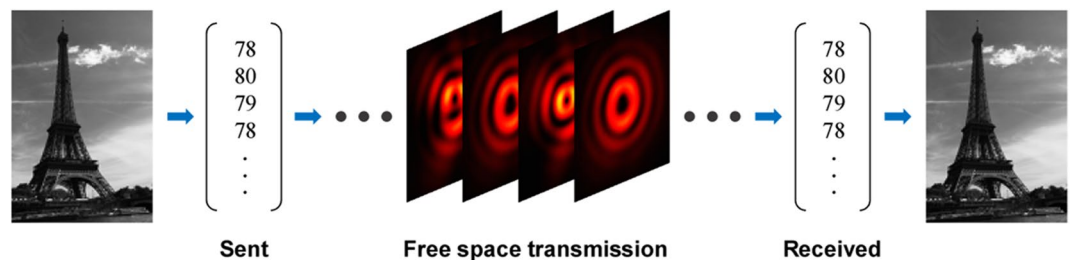
Figure 8 shows the one result of image transmission performed using the proposed optical link. As a test sample, an Eiffel Tower image of 100×150 pixels was sent. Using the proposed 8-bit encoding scheme, we were able to process a total of 120,000 binary data into 15,000 encoding symbols. The measured error rate was 0.05%, and the errors occurred only at 8 pixels: two due to the nearest OAM mode, eight due to the nearest radial mode. A correlation coefficient between the transmitted image and received image was higher than 0.99, which implies that the reconstructed image is almost identical to the transmitted one. Despite the successful image

Figure 7. (**a**) A part of the confusion matrix obtained with a test set and (**b**) a graph of test accuracy versus the applied phase shift. Displayed numbers from 64 to 95 correspond to spatial modes given by combination of a radial wavenumber 7.30 and OAM modes from 1.03 to 1.96. The values inside the matrix represent the number of recognized spatial mode. The diagonal elements of the matrix indicate that the transmitted mode is correctly recognized. A unit of the phase shift here is rad/π.

	0–31	32–63	64–95	96–127	128–159	160–191	192–223	224–255
k_r	7.10	7.20	7.30	7.40	7.50	7.60	7.70	7.80
l	1.03–1.96							

Table 2. 8-Bit encoding scheme.

Figure 8. Image transmission in the proposed optical link. An Eiffel Tower gray image of 100×150 pixels is transmitted and received. An error rate defined as the ratio of the number of incorrect pixels to the total number of pixels was 0.05%.

reconstruction with the very high correlation coefficient, a wrong prediction caused by the adjacent radial mode is not negligible. However, one can reduce the error through multiple measurements or proper post-processing algorithms[50]. For example, a simple way is to conduct image transmission multiple times and extract either the most frequently occurring number or average value per each pixel. The image was transmitted four times consecutively to investigate not only long-term stability but the error reduction, and the error rate of the reconstructed image by extracting the most frequent number per each pixel was 0.

Despite the impressive performance of the deep-learning-based optical link, it is necessary to discuss some critical issues for future practical applications. One thing is the transmission speed associated with data encoding and decoding. Encoding speed is determined by the frame rate of an SLM, which limits the symbol rate of the communication link. In the experiment, the hologram switching rate is set to 2.5 Hz, transmitting 20 bits of information per second. Even if we use the maximum rate of the SLM (60 Hz), the highest achievable data rate is only 480 bits per second. However, the demand for high-speed OAM switching can be addressed by using tunable integrated OAM devices[51,52] or DMD[53], which can increase the modulation rate up to a few tens of kilohertz or faster. Meanwhile, a camera with a frame rate faster than the refresh rate of a used modulator is needed to capture and decode laser spatial mode varying over time, and it could be addressed by using a CMOS camera. Assuming that the frame rate of a system, made up of a DMD and a CMOS camera, is 1 kHz, the transmission speed that can be achieved with 256-ary spatial mode encoding is just 8 (1×8) Kbit/s. However, it is possible to increase

the number of bits per symbol by employing more spatial modes and combining the OAM phase shift keying. Note that we achieved 8-bit encoding in an OAM subspace with the radial mode and deep-learning-based high-resolution recognition method. It indicates that the proposed link can transmit numerous data even if the size of the aperture restricts the range of available OAM modes. Besides, the capacity of the communication system can be increased further by combining other photonic DoF (wavelength and polarization)[10]. In other words, the speed can be doubled by combining two polarization states, and N times further by adding phase shift modulation of fractional modes. The maximum data speed is estimated to be a few hundred Kbit/s. Note that an optical link with a single wavelength was considered. This shows a relatively lower speed compared to fiber-optic communication systems, but the proposed OAM encoding scheme is able to improve physical layer security due to its inherent characteristic that does not depend on mathematical or quantum–mechanical encryption methods[5,54]. Therefore, the proposed optical link would also be applicable for military communication systems requiring high security.

For long-distance outdoor links, there exists atmospheric turbulence that distorts the structured phase and intensity and degrades the link performance. Methods of adaptive optics, based on a wavefront sensor[55] or deep-learning[37,38], can be taken into account to compensate for the deteriorating effects of turbulence. In the case of strong atmospheric turbulence, an active method based on ultrashort high-intensity laser filaments can produce a cleared optical channel by opto-mechanically expelling the droplets out of the beam area[56]. Besides, it is possible to achieve long-range propagation by means of nonlinear self-channeling of high-power laser pulses[57]. Compared to the ordinary spatial modes of an integer mode index, fractional modes are much sensitive to the external perturbation. We will investigate the transmission performance of the fractional mode encoding in different atmospheric turbulence levels and an effective method to mitigate intermodal crosstalk.

Conclusion

In conclusion, we experimentally demonstrated that both radial and azimuthal modes of BG beams, with the fractional mode spacing, can be used to transmit information over free space. 8-bit data are densely encoded in 256 spatial modes and successfully decoded by a deep-learning classifier. To achieve this, we first trained the designed neural network based on Alexnet architecture and tested its performance. The recognition accuracy of different fractional modes was nearly 100%, which demonstrates that the deep-learning decoder can simultaneously identify two independent spatial modes and accurately recognize extremely small differences between adjacent modes. Furthermore, the translational and rotational invariance of the trained CNN provided a reliable, stable model performance despite the sensitivity of fractional OAM beams to the optical alignment. Then, we transmitted a 100×150 grayscale image, a total of 120,000 bits, encoded with 15,000 data symbols by switching phase holograms, and the transmitted data was successfully recovered in the proposed optical link. In addition to the fractional modulation, the explicit phase shift applied to fractional OAM beams could provide an additional degree of freedom for information coding, which is achievable without degrading the performance of the proposed method. There are challenges to be overcome, such as modulation rate and atmospheric turbulence, but the proposed fractional mode encoding/deep-learning decoding scheme will provide an effective way to meet the growing demand for data traffic.

References

1. Gnauck, A. *et al.* Spectrally efficient long-haul WDM transmission using 224-Gb/s polarization-multiplexed 16-QAM. *J. Lightwave Technol.* **29**, 373–377 (2011).
2. Zhou, X. *et al.* 64-Tb/s, 8 b/s/Hz, PDM-36QAM transmission over 320 km using both pre- and post-transmission digital signal processing. *J. Lightwave Technol.* **29**, 571–577 (2011).
3. Sano, A. *et al.* Ultra-high capacity WDM transmission using spectrally-efficient PDM 16-QAM modulation and C- and extended L-band wideband optical amplification. *J. Lightwave Technol.* **29**, 578–586 (2011).
4. Ryf, R. *et al.* Mode-division multiplexing over 96 km of few-mode fiber using coherent 6 × 6 MIMO processing. *J. Lightwave Technol.* **30**, 521–531 (2012).
5. Gibson, G. *et al.* Free-space information transfer using light beams carrying orbital angular momentum. *Opt. Express* **12**, 5448–5456 (2004).
6. Čelechovský, R. & Bouchal, Z. Optical implementation of the vortex information channel. *New J. Phys.* **9**, 328–328 (2007).
7. Allen, L., Beijersbergen, M., Spreeuw, R. & Woerdman, J. Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes. *Phys. Rev. A* **45**, 8185–8189 (1992).
8. Allen, L., Padgett, M. & Babiker, M. The orbital angular momentum of light. *Prog. Opt.* https://doi.org/10.1016/s0079-6638(08)70391-3 (1999).
9. Du, J. & Wang, J. High-dimensional structured light coding/decoding for free-space optical communications free of obstructions. *Opt. Lett.* **40**, 4827–4830 (2015).
10. Trichili, A. *et al.* Optical communication beyond orbital angular momentum. *Sci. Rep.* **6**, 27674 (2016).
11. Trichili, A., Salem, A., Dudley, A., Zghal, M. & Forbes, A. Encoding information using Laguerre Gaussian modes over free space turbulence media. *Opt. Lett.* **41**, 3086–3089 (2016).
12. Kai, C., Huang, P., Shen, F., Zhou, H. & Guo, Z. Orbital angular momentum shift keying based optical communication system. *IEEE Photon. J.* **9**, 1–10 (2017).
13. Fu, S. *et al.* Experimental demonstration of free-space multi-state orbital angular momentum shift keying. *Opt. Express* **27**, 33111–33119 (2019).
14. Wang, J. *et al.* Terabit free-space data transmission employing orbital angular momentum multiplexing. *Nat. Photon.* **6**, 488–496 (2012).
15. Bozinovic, N. *et al.* Terabit-scale orbital angular momentum mode division multiplexing in fibers. *Science* **340**, 1545–1548 (2013).
16. Lei, T. *et al.* Massive individual orbital angular momentum channels for multiplexing enabled by Dammann gratings. *Light Sci. Appl.* **4**, e257 (2015).
17. Ahmed, N. *et al.* Mode-division-multiplexing of multiple Bessel-Gaussian beams carrying orbital-angular-momentum for obstruction-tolerant free-space optical and millimetre-wave communication links. *Sci. Rep.* **6**, 22082 (2016).

18. Zhu, L. *et al.* Orbital angular momentum mode groups multiplexing transmission over 2.6-km conventional multi-mode fiber. *Opt. Express* **25**, 25637–25645 (2017).
19. Zhu, L., Wang, A., Chen, S., Liu, J. & Wang, J. Orbital angular momentum mode multiplexed transmission in heterogeneous few-mode and multi-mode fiber network. *Opt. Lett.* **43**, 1894–1897 (2018).
20. Zhao, N., Li, X., Li, G. & Kahn, J. Capacity limits of spatially multiplexed free-space communication. *Nat. Photon.* **9**, 822–826 (2015).
21. Chen, M., Dholakia, K. & Mazilu, M. Is there an optimal basis to maximise optical information transfer?. *Sci. Rep.* **6**, 22821 (2016).
22. Götte, J., Franke-Arnold, S., Zambrini, R. & Barnett, S. Quantum formulation of fractional orbital angular momentum. *J. Mod. Opt.* **54**, 1723–1738 (2007).
23. Berry, M. Optical vortices evolving from helicoidal integer and fractional phase steps. *J. Opt. A: Pure Appl. Opt.* **6**, 259–268 (2004).
24. Götte, J. *et al.* Light beams with fractional orbital angular momentum and their vortex structure. *Opt. Express* **16**, 993–1006 (2008).
25. Zhang, N. *et al.* Analysis of fractional vortex beams using a vortex grating spectrum analyzer. *Appl. Opt.* **49**, 2456–2462 (2010).
26. Forbes, A., Dudley, A. & McLaren, M. Creation and detection of optical modes with spatial light modulators. *Adv. Opt. Photon.* **8**, 200–227 (2016).
27. Zheng, S. & Wang, J. Measuring orbital angular momentum (OAM) states of vortex beams with annular gratings. *Sci. Rep.* **7**, 40781 (2017).
28. Deng, D., Lin, M., Li, Y. & Zhao, H. Precision measurement of fractional orbital angular momentum. *Phys. Rev. Appl.* **12**, 014048 (2019).
29. Berkhout, G., Lavery, M., Padgett, M. & Beijersbergen, M. Measuring orbital angular momentum superpositions of light by mode transformation. *Opt. Lett.* **36**, 1863–1865 (2011).
30. Berger, B., Kahlert, M., Schmidt, D. & Assmann, M. Spectroscopy of fractional orbital angular momentum states. *Opt. Express* **26**, 32248–32258 (2018).
31. Xie, G. *et al.* Performance metrics and design considerations for a free-space optical orbital-angular-momentum–multiplexed communication link. *Optica* **2**, 357–365 (2015).
32. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
33. Balas, V., Roy, S., Sharma, D. & Samui, P. *Handbook of Deep Learning Applications* (Springer, New York, 2019).
34. Doster, T. & Watnik, A. Machine learning approach to OAM beam demultiplexing via convolutional neural networks. *Appl. Opt.* **56**, 3386–3396 (2017).
35. Lohani, S., Knutson, E., O'Donnell, M., Huver, S. & Glasser, R. On the use of deep neural networks in optical communications. *Appl. Opt.* **57**, 4180–4190 (2018).
36. Li, J., Zhang, M., Wang, D., Wu, S. & Zhan, Y. Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication. *Opt. Express* **26**, 10494–10508 (2018).
37. Lohani, S. & Glasser, R. Turbulence correction with artificial neural networks. *Opt. Lett.* **43**, 2611–2614 (2018).
38. Liu, J. *et al.* Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication. *Opt Express* **27**, 16671–16688 (2019).
39. Mao, Z. *et al.* Broad bandwidth and highly efficient recognition of optical vortex modes achieved by the neural-network approach. *Phys. Rev. Appl.* **13**, 034063 (2020).
40. Eric, K. Quantifying Translation-Invariance in Convolutional Neural Networks. Preprint at https://arxiv.org/abs/1801.01450 (2017).
41. Liu, Z., Yan, S., Liu, H. & Chen, X. Superhigh-resolution recognition of optical vortex modes assisted by a deep-learning method. *Phys. Rev. Lett.* **123**, 183902 (2019).
42. Paterson, C. & Smith, R. Higher-order Bessel waves produced by axicon-type computer-generated holograms. *Opt. Commun.* **124**, 121–130 (1996).
43. Yang, Z., Zhang, X., Bai, C. & Wang, M. Nondiffracting light beams carrying fractional orbital angular momentum. *J. Opt. Soc. Am. A* **35**, 452–461 (2018).
44. He, K., Zhang, X., Ren, S. & Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *In Proc. IEEE International Conference on Computer Vision* 1026–1034 (2015).
45. Srivastava, N., Hinton, G., Alex, K., Ilya, S. & Ruslan, S. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
46. Kingma, D. & Ba, J. Adam: A Method for Stochastic Optimization. Preprint at https://arxiv.org/abs/1412.6980 (2014).
47. Ren, Y. *et al.* Experimental generation of Laguerre-Gaussian beam using digital micromirror device. *Appl. Opt.* **49**, 1838–1844 (2010).
48. Ng, A. Feature selection, L1 vs. L2 regularization, and rotational invariance. *In Proc. 21st International Conference on Machine Learning* 78–85 (2004).
49. Ioffe, S. & Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. Preprint at https://arxiv.org/abs/1502.03167 (2015).
50. Han, Y. & Leou, J. Detection and correction of transmission errors in JPEG images. *IEEE Trans. Circuits Syst. Video Technol.* **8**, 221–231 (1998).
51. Strain, M. *et al.* Fast electrical switching of orbital angular momentum modes using ultra-compact integrated vortex emitters. *Nat. Commun.* **5**, 4856 (2014).
52. Xie, Z. *et al.* Ultra-broadband on-chip twisted light emitter for optical communications. *Light Sci. Appl.* **7**, 18001 (2018).
53. Hu, X. *et al.* Dynamic shaping of orbital-angular-momentum beams for information encoding. *Opt. Express* **26**, 1796–1808 (2018).
54. Wang, T., Gariano, J. & Djordjevic, I. Employing Bessel-Gaussian beams to improve physical-layer security in free-space optical communications. *IEEE Photon. J.* **10**, 1–13 (2018).
55. Ren, Y. *et al.* Adaptive optics compensation of multiple orbital angular momentum beams propagating through emulated atmospheric turbulence. *Opt. Lett.* **39**, 2845–2848 (2014).
56. Schimmel, G., Produit, T., Mongin, D., Kasparian, J. & Wolf, J. Free space laser telecommunication through fog. *Optica* **5**, 1338–1341 (2018).
57. Peñano, J., Palastro, J., Hafizi, B., Helle, M. & DiComo, G. Self-channeling of high-power laser pulses through strong atmospheric turbulence. *Phys. Rev. A* **96**, 013829 (2017).

Acknowledgements

Author contributions
Y.N. performed the experiments and analyzed data under the guidance of DK.K.; DK.K. supervised the experiment and the project; all the authors contributed to writing the manuscript.

Competing interests
The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.-K.K.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.