Deep Learning with Persistent Homology for Orbital Angular Momentum (OAM) Decoding

Soheil Rostami, Walid Saad, and Choong Seon Hong

Abstract—Orbital angular momentum (OAM)-encoding has recently emerged as an effective approach for increasing the channel capacity of free-space optical communications. In this paper, OAM-based decoding is formulated as a supervised classification problem. To maintain lower error rate in presence of severe atmospheric turbulence, a new approach that combines effective machine learning tools from persistent homology and convolutional neural networks (CNNs) is proposed to decode the OAM modes. A Gaussian kernel with learnable parameters is proposed in order to connect persistent homology to CNN, allowing the system to extract and distinguish robust and unique topological features for the OAM modes. Simulation results show that the proposed approach achieves up to 20% gains in classification accuracy rate over state-of-the-art of method based on only CNNs. These results essentially show that geometric and topological features play a pivotal role in the OAM mode classification problem.

Index Terms—OAM, convolutional neural networks, persistent homology, free-space optical communication.

I. INTRODUCTION

Free-space optical (FSO) communication is an effective approach for fixed point-to-point communication, such as backhaul connectivity and fiber backup over distances up to several kilometers [1]. In order to increase the transmission capacity of FSO communication, space division multiplexing can be exploited. In particular, the use of the orbital angular momentum (OAM) of a light beam has been recently proposed to realize space division multiplexing in FSO communication systems [2]. OAM is expected to play a major role in many emerging communication systems [3].

Theoretically, Laguerre-Gaussian (LG) beams can carry an infinite values of OAM modes, by encoding a bit-tuple of information as superposition of OAM modes [2]. Therefore, the transmission capacity can be increased with orders-of-magnitude by allowing beams with different modes to be multiplexed together and transmitted over the same communication link. However, in FSO communications, the high sensitivity of the spatial structure of a light beam to atmospheric conditions such as turbulence can cause cross-talk among adjacent OAM modes. This, in turn, makes it challenging to perform error-free OAM mode detection particularly when a large number of OAM modes are used [2].

S. Rostami is with Huawei Technologies Oy (Finland) Co. Ltd, Helsinki, Finland. E-mail: soheil.rostami1@huawei.com.

W. Saad is with the Bradley Department of Electrical and Computer Engineering, Virginia Tech, USA. E-mail: walids@vt.edu.

C. S. Hong is with Department of Computer Science and Engineering, Kyung Hee University, South Korea. E-mail: cshong@khu.ac.kr.

Conventionally, an optical solution based on coherent detection, known as conjugate-mode sorting method is applied for OAM detection as done in [4]. For the coherent detection of OAM modes, both the transmitter and receiver can use spatial light modulators (SLMs). In [4], a first proof-of-concept experiment, based on coherent detection, is developed to utilize OAM modes by defining an unlimited dimensional discrete Hilbert space. Recently, neural network approaches have been adopted in [5] and [6] in order to enhance OAM decoding by relying only on an intensity image of the unique multiplexing patterns. The authors in [5] proposed a deep neural network approach capable of simultaneously differentiating 110 OAM modes with classification error rate of less than 30%. Meanwhile, the authors in [6], used convolutional neural networks (CNN) to differentiate 32 OAM modes with over 99% accuracy under high levels of turbulence. The work in [6] also showed that this new method is robust to various environmental parameters. However, the performance of the solutions proposed in [5] and [6] degrades significantly for large number of OAM modes and high levels of turbulence, thus motivating the need for new, turbulence-robust solutions.

Recently, OAM communications has attracted significant attention in the wireless communication literature. For instance, in [7], the authors study the problem of enhancing spectrum efficiency for multi-user access with different OAM modes for two-tier wireless networks. Meanwhile, in [8], the authors conducted important experiments to characterize the OAM phase properties for long-distance transmission.

The main contribution of this paper is a novel framework that exploits the powerful machine learning tools of *persistent homology* [9] to enhance OAM mode detection, in presence of turbulence. In particular, the proposed approach combines a persistent homology-based input layer to a CNN-based OAM decoder. This input layer allows inference of topological and geometrical information rather than the intensity images, for OAM mode detection. Our results show that exploiting topological features is more effective than phase fronts, applied in intensity images. In particular, simulation results show that the proposed approach outperforms CNN, in terms of the classification accuracy rate, by 10% in presence of severe atmospheric turbulence and a large number of OAM modes.

II. SYSTEM MODEL

We consider an OAM communication system composed of a single transmitter and receiver pair. The transmitter communicates one out of M possible n-bit length messages $s \in \mathcal{M}$, where $M = 2^n$. To transmit each message, n

different OAM modes $\{c_1, ..., c_n\}$ are superpositioned[1], and the corresponding transmitted beam $\boldsymbol{x} \in \mathbb{C}^n$ is sent over the channel. Hence, the transmitter can be seen as a mapping function, $\mathfrak{T} : \mathcal{M} \mapsto \mathbb{C}^n$. At the receiver, the received beam $\boldsymbol{y} \in \mathbb{C}^n$ is a noisy and distorted version of the transmitted beam. The receiver must produce an estimate $\hat{s}$ of the original message s, and, hence, the receiver can be seen as a mapping function, $\mathfrak{R} : \mathbb{C}^n \mapsto \mathcal{M}$. Both $\boldsymbol{x}$ and $\boldsymbol{y}$ are vectors with each element corresponding, respectively, to the intensity level of the transmitted or received beam's pixel. The length of $\boldsymbol{x}$ and $\boldsymbol{y}$ equals to the number of pixels per beam p, which depends on the pixel size of a charge-coupled device image sensor.

The considered OAM communication system must reduce the message error rate ($P_e = \Pr(\hat{s} \neq s)$) for a high number of OAM modes and severe turbulence levels. Moreover, some OAM modes are more robust to atmospheric turbulence, and, thus, the n OAM modes that are used for encoding a message can be selected in such a way to have the least sensitivity to the channel conditions. Here, we assume that the set of OAM modes, i.e. $\{c_1, ..., c_n\}$ is fixed.

A CNN learning technique can be used for OAM decoding with the objective of minimizing the message error rate. Such a neural network can be modeled as a function composition chain of functions, $f(\boldsymbol{y}) = a_L \circ q_L \circ a_{L-1} \circ q_{L-1} \circ \cdots \circ a_1 \circ q_1(\boldsymbol{y})$; where a_j and q_j are convolutional (conv1,...,convL) and pooling (pool1,...,poolL) functions, respectively; and L is the number of layers in the network. The convolution filters apply convolution operations to the input, and then pooling applies a moving window to choose the maximum value over a local region. The final layer (based on softmax activation) of the network represents the M messages. Once trained, during the testing phase, the CNN yields a set of probabilities $\{f_1, ..., f_m, ..., f_M\}$, where f_m is essentially the probability that the CNN's input message is classified as message m. Message i is detected as the received message if $f_m < f_i$ for all $m \in \mathcal{M}$ and $m \neq i$.

Although CNN can be effective for improving OAM reception as shown in [6], relying solely on CNN for the purpose of OAM detection can fail if the number of OAM modes is large. In particular, CNN performs its decoding by relying solely on the intensity profiles while ignoring the topological structure of each OAM mode. However, in practice, each OAM mode, also known as a topological charge, possesses a unique twisted structure. Such topological structures contain more discriminate features than intensity profiles. Hence, in order to enhance the effectiveness of OAM decoding, it is desirable to design new learning techniques that can exploit such topological structures, which has not been done in prior works [5], [6]. In this context, persistent homology has recently emerged as a powerful tool for addressing the problem of topological feature detection and shape recognition. By means of persistent homology, a shape is represented with a family of simplicial complexes, indexed by a proximity parameter

---

¹The throughput of the FSO communication link is linearly dependent on n, and hence it is beneficial to increase n as much as possible to achieve higher throughput for a given error rate.

TABLE I
SUMMARY OF MAIN NOTATIONS.

Variable	Definition
$\boldsymbol{\mu_i} = [\mu_{i,1}, \mu_{i,2}]^T$	Location parameters of Gaussian Kernel
σ_i	Standard deviation of Gaussian Kernel
ν	Parameter to control the effect of the persistence and computational complexity
$\mathcal{B}$	Persistence diagram (can be visualized as multisets of intervals)
$\|.\|$	Cardinality of a set

[9]. This converts different shapes into global topological objects. The output of persistent homology is in the form of a parameterized version of a Betti number, and can be illustrated by a persistence diagram [9]. Depending on the combination of OAM modes used in the transmitted beam, the received beam will exhibit a unique topological structure which differentiates it from other feasible combinations of OAM modes. Therefore, persistent homology can be used to extract such features which can, in turn, be exploited by the CNN to classify the received message. Table I summarizes our notations.

Due to space limitation, the mathematical intricacies of persistent homology are not provided in detail; we refer the reader to [9] for algebraic-topological aspects of persistent homology. Next, we propose a novel framework that employs persistent homology as an input layer for CNN in order to probe topological features of OAM modes and enhance the OAM receiver performance.

III. PROPOSED FRAMEWORK

The utilization of unique topological structures of OAM modes is promising to enhance OAM mode detection because the topological features are independent of any particular coordinate. As result, no optical alignment process is needed. Furthermore, the topological features of OAM beams are deformation invariant, i.e. their topology does not change if the shape is stretched or compressed, due to atmospheric turbulence. Also, topological analysis of OAM beams enhances robustness to noise, enabling long distance communication.

Due to the superposition of n-orthogonal topological charges, each transmitted beam $\boldsymbol{x}$ and hence corresponding received beam $\boldsymbol{y}$ will exhibit a unique topological structure. In this regard, persistent homology will allow us to replace each point of an OAM topological structure with a solid sphere and, then, record the topological features of the union of these solid spheres as a function of their radius. When the radius increases, new components can be created, or existing components can merge together, or can be eliminated. The 0th order topological feature captures the connected components, the 1st order topological feature captures regions forming a loop structure, and 2nd order topological feature captures regions forming a void structure. The radius at which the jth component, referred to as p_j, is created is called its *birth time* b_j and the radius at which component p_j is eliminated, or merged with another component that has an earlier birth time, is called its *death time* d_j. For our OAM system,
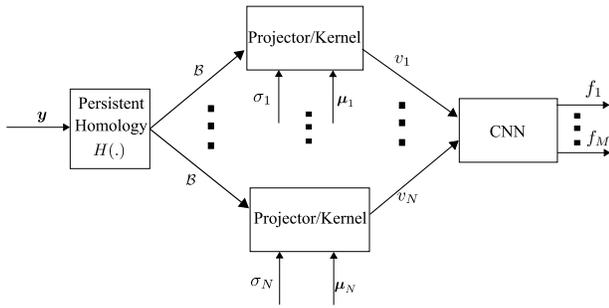
Figure 1. The overall functionality of the proposed method used by an OAM receiver to classify OAM modes, both projector and CNN are trained as a single neural network.

Figure 2. The transmitted OAM beam when $M = 16$.

persistent homology detects the connected components (0th order), loops (1st order), and voids (2nd order) of the received beam. Meanwhile, the corresponding (b_j, d_j)-tuples measure how persistent the received beam's topological features are. In other words, the difference $(d_j - b_j)$ is known as the *lifetime* of the topological feature and measures its robustness and prominence.

The persistence diagram can be visualized as multisets of intervals $\mathcal{B} = \{p_j = (b_j, d_j) : 0 < b_j < d_j, j = 1, ..., |\mathcal{B}|\}$. We define a *persistent homology* as a mapping function $\mathfrak{H}$, which maps the received beam $\boldsymbol{y}$ to its corresponding persistence diagram $\mathcal{B}$, i.e. $\boldsymbol{y} \overset{\mathfrak{H}}{\mapsto} \mathcal{B}$. The persistence diagram $\mathcal{B}$ does not possess a Hilbert space structure due to its unusual structure as multisets. As result, machine learning algorithms (e.g. CNN) operating on a Hilbert space cannot be applied to persistence diagrams. However, a persistence diagram can embed the set of persistence diagrams into a Hilbert space by applying *kernels* to map the topological features into machine learning compatible representation [10]. However, such kernels are pre-defined and fixed, and therefore agnostic to the learning stage of the supervised classification. For our case, we propose a modified version of a Gaussian kernel with learnable parameters. In particular, the parameters of our kernel are tuned using back-propagation.

For the CNN of our OAM decoder, we introduce a persistent homology-based input layer that is defined by projecting any component p_j with respect to a collection of kernels. Therefore, each component p_j, is transformed to a single value z_j (i.e. $p_j \overset{G_{\sigma_i, \mu_i}}{\mapsto} z_j$) as follows, $G_{\sigma_i, \mu_i}(p_j) = e^{-\frac{\|p_j - \mu_i\|}{2\sigma_i^2}}$, where $\mu_i = [\mu_{i,1}, \mu_{i,2}]^T$ is a location parameter and $\sigma_i > 0$ is standard deviation. Both are learned using a backpropagation learning method during the training stage.

Due to fact that components close to the diagonal of persistence diagram are mainly noise, we can remove these components in our analysis, i.e. $z_j = 0$ if $d_j - b_j < \nu$ where ν is a constant. Consequently, not only the contribution of noise is discounted but also computational complexity is reduced.

The persistence diagram $\mathcal{B}$ per kernel is projected into a single value $v_i = \sum_{j=1}^{|\mathcal{B}|} G_{\sigma_i, \mu_i}(p_j)$. By projecting $\mathcal{B}$ with respect to N kernels with different parameters, the *topological feature vector* $\boldsymbol{v}$ of the received OAM beam $\boldsymbol{y}$ can be constructed, i.e. $z_j \overset{\mathfrak{v}}{\mapsto} \boldsymbol{v}$, where $\boldsymbol{v} = [v_1, ..., v_N]^T$.
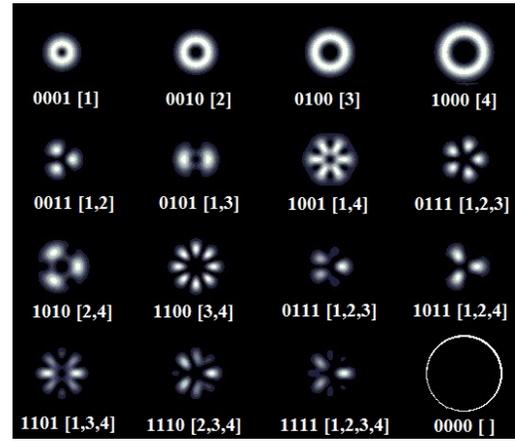
The end-to-end operation of the framework is shown in Fig. 1. First, $\mathcal{B}$ is computed using any standard persistent homology algorithm (e.g. [9]) over the received OAM beam $\boldsymbol{y}$, and then $\mathcal{B}$ is projected to $\boldsymbol{v}$ which itself is an input layer for the CNN. Therefore, the overall function of the resulting neural network can be written as follows,

$$f(\boldsymbol{y}) = a_L \circ q_L \circ a_{L-1} \circ q_{L-1} \circ \cdots \circ a_1 \circ q_1 \circ \mathfrak{v} \circ \mathfrak{H}(\boldsymbol{y}), \quad (1)$$

where σ_i and μ_i (for all $i \in \{1, ..., N\}$) and corresponding parameters of L layers are learned during training.

An example of a computer-generated phase diagram without turbulence for superposition of an OAM mode set $\{1, 2, 3, 4\}$ is shown in Fig. 2. The caption of each sub-image is the encoded 4 bit-length message and the modes that are active (set of integers in brackets). For instance, $0101[1, 3]$ means that for four-bit message 0101, OAM modes 1 and 3 are superpositioned and transmitted. Now, based on Fig. 2, we can illustrate the reasons why the proposed approach, that combines persistent homology and CNN, is more effective than CNN alone, for OAM communications. In this regard, the received beam of OAM communication has a unique "twisted" structure (e.g. see Fig. 2). The persistent homology acts as a topological feature selector which can automatically choose those topological OAM features that contribute the most to the output. However, without persistent homology, the trained CNN model will rely on the irrelevant features and noise, which can then reduce the accuracy of the OAM decoder. Furthermore, by using persistent homology, each received OAM beam can be properly categorized to its unique topological structures through topological feature vector $\boldsymbol{v}$. Last but not least, the proposed method reduces training time and algorithm complexity by removing the irrelevant features and the components that most likely are linked to noise.

In order to intuitively explain how the proposed method outperforms conventional CNN, the persistence diagram, phase front and topological structure of a pair of 8-bit length OAM messages both in transmit and receive are provided in Fig. 3. In particular, two OAM beams corresponding to the 82nd and 91st messages are chosen. The main reason for choosing

Table II
PARAMETER ANALYSIS OF LAYERS USED IN THE PROPOSED METHOD
(PH+CNN) AND BASELINE (CNN: CONV1, POOL1,..., CONV5, POOL5,
FC1, FC2, FC3) WHEN $N = 1000$.

Proposed method	Baseline	Layer	Kernel	Parameters	FLOPs
PH	-	PH	1×1	3 K	1 G
CNN	CNN	conv1	11×11	35 K	40 G
		pool1	3×3	0	1 G
		conv2	5×5	615 K	120 G
		pool2	3×3	0	1 G
		conv3	3×3	885 K	40 G
		conv4	3×3	1.33 M	60 G
		conv5	3×3	885 K	60 G
		pool5	3×3	0	1 G
		fc6	1×1	38 M	10 G
		fc7	1×1	17 M	5 G
		fc8	1×1	4 M	2 G

Table III
OVERALL FLOP COUNTS, NUMBER OF PARAMETER AND RUNNING TIME
OF PROPOSED METHOD (PH+CNN) AND BASELINE (CNN) WHEN
$N = 1000$.

Method	Parameters	FLOPs	Running Time
PH+CNN	62.753 M	341 G	42.1 ms
CNN	62.75 M	340 G	42 ms

the pair as example is that, CNN frequently misclassifies 82nd message with 91st. In contrast, our proposed method classifies this particular example correctly most of time. It is clear that for both received messages, their persistence diagrams, especially the 2nd order topological features (shown in black) do not vary much compared to the phase diagrams. As result, the CNN which relies on phase diagrams, will misclassify the received OAM beam. Significant features of 3D topological shape (shown in blue) are far from the diagonal in the persistence diagram, and their corresponding points in the diagram do not change significantly. Furthermore, Fig. 3 e) illustrates the output layer (the probabilistic values between 0 and 1) resulting from the proposed framework and conventional CNN, when the transmit message's index is 82. Clearly, in the proposed framework, f_{82} is much higher than other f_m for all $1 \leq m \leq 128$ and $m \neq 82$. However, in case of conventional CNN, f_{91} is higher than other f_m for all $1 \leq m \leq 128$ and $m \neq 91$, and thus the CNN misclassifies the 82nd message as the 91st.

Our network is based on Alexnet [11] and is composed of 5 conv-pooling layers (explained in Section II) and 3 fully connected (fc) layers. The detailed network architectures of the proposed method (PH, conv1, pool1,..., conv5, pool5, fc1, fc2, fc3) and baseline (conv1, pool1,..., conv5, pool5, fc1, fc2, fc3) are shown in Table II. In order to remove the dependency of the algorithm's running time on hardware, the floating-point instruction (FLOP) counts are calculated. The FLOP counts of backward propagation layers are listed in Table II, from which we can see that processing overhead of computing persistent homology (PH) is much lower than the rest of the convolutional layers. Also, Table III shows that the running time and FLOP counts of the proposed method (including persistent homology plus Alexnet-based CNN) are just 0.3% and 0.5% higher than the baseline (Alexnet) when both are running on GTX1080, respectively.
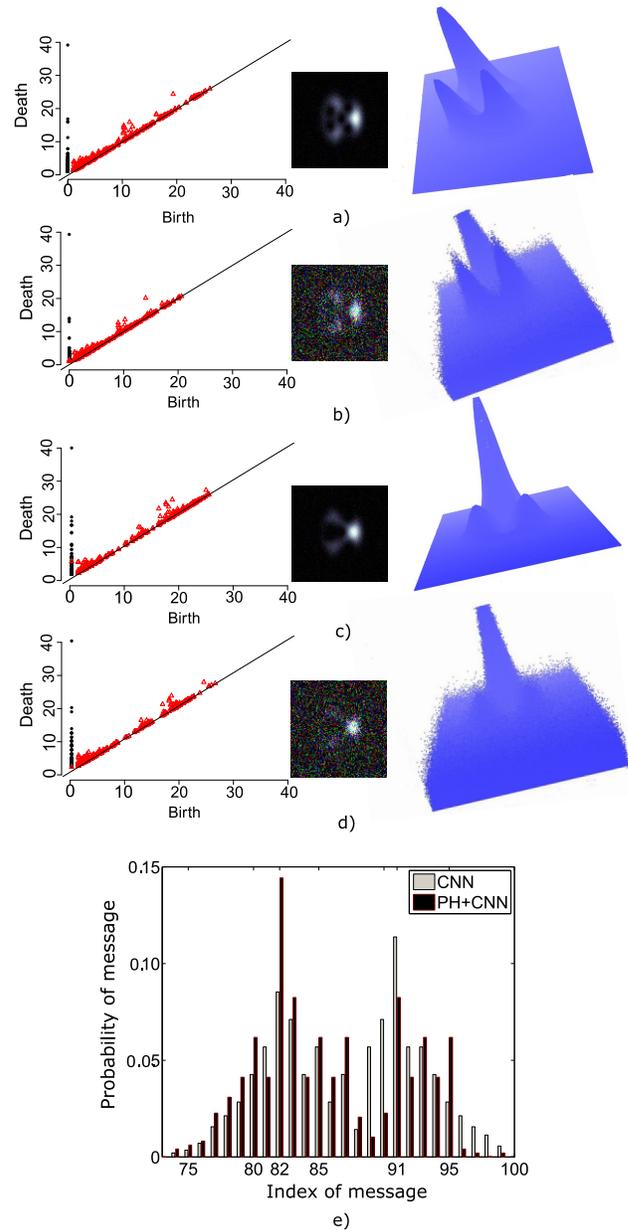

Figure 3. The OAM decoding of $M = 128$; due to importance, the 1st and 2nd order topological feature are shown with black and red, respectively; a) the 82nd message is shown at the transmit side ($\boldsymbol{x}$); b) the 82^{ed} message is shown at the receive side ($\boldsymbol{y}$); c) the 91st message is shown at the transmit side ($\boldsymbol{x}$); d) the 91st message is shown at the receive side ($\boldsymbol{y}$); e) f_m is illustrated for both proposed method PH+CNN and CNN.

IV. SIMULATION RESULTS AND ANALYSIS

In this section, we demonstrate the performance of the proposed method in terms of accuracy ($\Pr(\hat{s} = s) = 1 - P_e$), and we compare it with the CNN-based method (Alexnet) of [6]. For the purpose of training and testing the framework, the first-n-adjacent OAM modes, ($n \in \{6, ..., 16\}$) are generated and superpositioned, numerically.

Additionally, the effect of varying the turbulence level and the message length (which is equal to the number of OAM modes) on the accuracy rate $(1 - P_e)$ are investigated. Turbulence causes a random phase along the propagation path

of the beam [12], and the turbulence level, denoted by T, can be quantified as the ratio between linear dimension of the SLM and the Fried's parameter [12]. We compare the performance of the proposed method, called "PH+CNN", with those achieved by the CNN in [6], and we call it "CNN".

Our dataset is generated numerically by using the "basic paraxial optics toolkit" in MATLAB [13], and it consists of a balanced group size (per OAM) of 10000 phase diagrams that contains a variable amount of turbulence. To train the CNN, we split the data collected into two separate sets a training set (85%) and testing set (15%) of the overall dataset. The training and testing sets are completely independent of one another and do not share any turbulence realizations. The training and testing set for the PH+CNN use the same turbulence realizations as those used for the CNN. In our simulations, the most meaningful results were found by empirically setting ν to 0.1. Therefore, all results are reported for $\nu = 0.1$. By adjusting ν, we can control and discount the effect of the points with low persistence and computational complexity (explained in Section III). Furthermore, for computing persistent homology, the "R package TDA" [14] is used.

Fig. 4 shows the accuracy of decoding for different turbulence levels and bit lengths of CNN and PH+CNN. Fig. 4 shows that, PH+CNN outperforms CNN for different turbulence levels and bit lengths. Fig. 4 shows that the performance of both methods degrades rapidly with increasing turbulence level and bit length. However, for higher turbulence levels, the performance of PH+CNN is much better than CNN while they perform similarly at low turbulence levels. It can be seen that for a severe turbulence level of 21, for a message with 16-bit length, PH+CNN yields up to 20% better accuracy than CNN. This demonstrates that the PH+CNN method can allow a larger message size or OAM number and can increase the capacity further even for high turbulence levels.

V. CONCLUSIONS

In this paper, we have proposed a new approach that combines persistent homology and CNN to decode OAM modes. We have shown that, after training the proposed method on set of each superpositioned OAM mode, a high accuracy, rate even with large number of OAM modes, in presence of severe atmospheric turbulence can be achieved. Numerical results have shown a substantial increase in the number of simultaneously-discriminated OAM modes under turbulence channel with better accuracy than CNN. Future work can consider the design of an autoencoder to select OAM modes with highest discriminative topological features as the messages, for given channel conditions, data rates and error rate requirements, adaptively.

REFERENCES

[1] M. A. Khalighi and M. Uysal, "Survey on free space optical communication: A communication theory perspective," *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2231–2258, Fourth Quarter 2014.
[2] J. Wang, J.-Y. Yang, I. M. Fazal, N. Ahmed, Y. Yan, H. Huang, Y. Ren, Y. Yue, S. Dolinar, M. Tur, and A. E. Willner, "Terabit free-space data transmission employing orbital angular momentum multiplexing," *Nature Photonics*, vol. 6, no. 7, p. 488, June 2012.
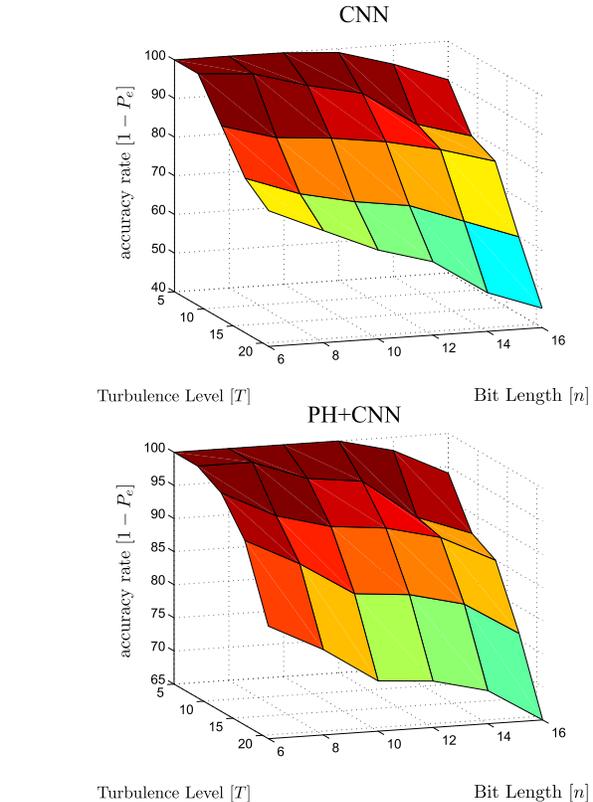
Figure 4. Decoding accuracy of CNN (top) and PH+CNN (bottom) as a function of bit lengths and turbulence levels.

[3] W. Saad, M. Bennis, and M. Chen, "A vision of 6G wireless systems: applications, trends, technologies, and open research problems," *IEEE Network*, to appear, 2019.
[4] A. Mair, A. Vaziri, G. Weihs, and A. Zeilinger, "Entanglement of the orbital angular momentum states of photons," *Nature*, vol. 412, no. 6844, p. 313, July 2001.
[5] E. M. 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 *Optical Engineering + Applications*, vol. 9970, Sep 2016, p. 997013.
[6] T. Doster and A. T. Watnik, "Machine learning approach to OAM beam demultiplexing via convolutional neural networks," *Applied Optics*, vol. 56, no. 12, pp. 3386–3396, 2017.
[7] W. Cheng, W. Zhang, H. Jing, S. Gao, and H. Zhang, "Orbital angular momentum for wireless communications," *IEEE Wireless Communications*, vol. 26, no. 1, pp. 100–107, February 2019.
[8] Y. Yao, X. Liang, W. Zhu, J. Geng and R. Jin, "Experiments of Orbital Angular Momentum Phase Properties for Long-Distance Transmission," *IEEE Access*, vol. 7, pp. 62 689–62 694, 2019.
[9] A. J. Zomorodian, *Topology for Computing*, ser. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2005.
[10] G. Kusano, Y. Hiraoka, and K. Fukumizu, "Persistence weighted gaussian kernel for topological data analysis," in *Proc. ICML'16*, NY, USA, June 2016, pp. 2004–2013.
[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. NIPS 2012*. Curran Associates, Inc., 2012, pp. 1097–1105.
[12] L. Andrews, "An analytical model for the refractive index power spectrum and its application to optical scintillations in the atmosphere," *Journal of Modern Optics*, vol. 39, no. 9, pp. 1849–1853, 1992.
[13] A. M. Gretarsson, "Basic paraxial optics toolkit—MATLAB central."
[14] B. T. Fasy, J. Kim, F. Lecci, and C. Maria, "Introduction to the R package TDA," *arXiv preprint arXiv:1411.1830*, Nov 2014.