

Received April 8, 2021, accepted May 6, 2021, date of publication May 25, 2021, date of current version August 12, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3083524

Performance Analysis of 3D Video Transmission Over Deep-Learning-Based Multi-Coded N-ary Orbital Angular Momentum FSO System

SHIMAA A. EL-MEADAWY¹, HOSSAM M. H. SHALABY^{2,3}, (Senior Member, IEEE),
NABIL A. ISMAIL⁴, AHMED E. A. FARGHAL⁵, FATHI E. ABD EL-SAMIE^{1,6},
MOHAMMED ABD-ELNABY⁷, AND WALID EL-SHAFAI^{1,8}

¹Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

²Department of Electrical Engineering, Faculty of Engineering, Alexandria University, Alexandria 21544, Egypt

³Department of Electronics and Communication, Egypt-Japan University of Science and Technology, Alexandria 21934, Egypt

⁴Department of Computer Science and Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

⁵Department of Electrical Engineering, Faculty of Engineering, Sohag University, Sohag 82524, Egypt

⁶Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia

⁷Department of Computer Engineering, College of Computers and Information Technology, Taif University, Taif 21944, Saudi Arabia

⁸Security Engineering Laboratory, Computer Science Department, Prince Sultan University, Riyadh 11586, Saudi Arabia

Corresponding author: Walid El-Shafai (eng.waled.elshafai@gmail.com)

This work was supported by Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif University, Taif, Saudi Arabia.

ABSTRACT Orbital angular momentum-shift keying (OAM-SK), which is the rapid switching of OAM modes, is vital but seriously impeded by the deficiency of OAM demodulation techniques, particularly when videos are transmitted over the system. Thus, in this paper, 3D chaotic interleaved multi-coded video frames (VFs) are conveyed via an N-OAM-SK free-space optical (FSO) communication system to enhance the reliability and efficiency of video communication. To tackle the defects of the OAM-SK-FSO mechanism, two efficient deep learning (DL) techniques, namely convolution recurrent neural network (CRNN) and 3D convolution neural network (3DCNN) are used to decode OAM modes with a low bit error rate (BER). Moreover, a graphics processing unit (GPU) is used to accelerate the training process with slight power consumption. The utilized datasets for OAM states are generated by applying different scenarios using a trial-and-error method. The simulation results imply that LDPC-coded VFs achieve the largest peak signal-to-noise ratios (PSNRs) and the lowest BERs using the 16-OAM-SK model. Both 3DCNN and CRNN techniques have nearly the same performance, but this performance deteriorates in the case of larger dataset classes. Moreover, the GPU accelerates the performance by almost 67.6% and 36.9% for the CRNN and 3DCNN techniques, respectively. These two DL techniques are more effective in evaluating the classification accuracy than the other traditional techniques by almost 10 – 20%.

INDEX TERMS 3DCNN, chaotic interleaver, OAM-SK, CRNN, turbulence channel.

I. INTRODUCTION

In recent years, free space optic (FSO) communication, which offers a straightforward tool to transfer excessive bit rates, has gained more and more interest as a complement or an alternative to the line-of-sight (LoS) radio relay link-age communication [1]. Recently, FSO using orbital angular momentum (OAM) has attracted a considerable attention

The associate editor coordinating the review of this manuscript and approving it for publication was Rui Wang¹.

as an emerging candidate for technology beyond the 5th generation, due to its potential as a means of permitting high-speed wireless transmission [2]. Since OAM can be used to establish multiple independent channels, wireless OAM multiplexing can effectively increase the transmission rate in a point-to-point link, such as wireless back-haul and/or fronthaul [3]. In [4], Kai *et al.* suggested an FSO communication scheme that uses OAM-shift keying (OAM-SK) as a modulation technique, where an encoded image is transmitted by a series of OAM states. The recov-

ered image displays high fidelity compared to the original image with a novel phase hologram after free-space transmission. In [5], the M -ary OAM-SK modulation efficiency was investigated with a Laguerre-Gauss (LG) beam considered over ocean turbulence. The results show that the optimum transmitted OAM mode set is mainly limited by the interfering energy and the decaying productive energy. In [6], it was demonstrated that the direct demodulation approach provides a feasible way to demodulate OAM-SK signals and can demonstrate potential in complex OAM modulation communication.

The three complete families of exact and orthogonal solutions for the paraxial wave equation are Hermite-Gaussian modes (HGMs), Laguerre-Gaussian modes (LGMs), and Ince-Gaussian modes (IGMs) [7]. Thirugnanasambandam *et al.* discovered a way to generate very-high order LGMs in end-pumped Yb: YAG ceramic lasers. Chen *et al.* proposed an approach to generate LGMs in end-pumped solid-state lasers (SSLs). Researchers recently invented a method for selectively producing a variety of higher-order LGMs from end-pumped SSLs by moving the pump to one of the brightest spots (target spot) from which the stimulated emission is generated. In the literature, a crosshair with a laterally uniform gain distribution is often introduced into the laser cavity in experiments involving higher-order HGMs in gas lasers. The LGM approaches HGMs as the ellipticity of the LGM approaches infinity [7].

The LG laser mode is the most common type of a helically-phased beam. A mode function is a mathematical expression that describes the magnitude and phase of the electric field at various points in the cross section. This function is the product of a Hermite polynomial and a Gaussian function for most laser beams without helical phasing. The HGMs have several intensity maxima arrayed in a rectilinear pattern and separated by intensity zeros, depending on the order of the polynomials. Due to the existence of an explicit phase factor in cylindrical LGMs [8], they are the natural choice for representing OAM rays. Although LGMs are produced directly in laser systems, they are more easily produced by the conversion of HG beams using cylindrical lenses. An analogous trick was introduced by the Leiden group employing cylindrical lenses to transform an HG beam without angular momentum into an LG beam with OAM. Although the conversion process is very successful, each LGM requires a specific HGM to begin with. This restriction limits the number of LGMs that can be generated. As a result, numerically computed holograms have become the most popular tool for generating helical beams. From the same initial beam, such holograms produce beams with any desired OAM mode [8].

Due to the merits of the 3D multiview video coding (MVV) applications in future multimedia communication, it will be recommended for wireless communication networks [9]. For active transmission of 3D MVV, the improvement of encoding performance is achieved by taking into consideration the sequential and longitudinal correlations amongst frames in

the same video and the merit of the inter-view matching inside different streams [10].

Recently, the absenteeism of twirls in convolution neural networks (CNNs), and feed-forward neural networks (FFNNs), has made them insufficient techniques for use in sequence labeling [11]. Recurrent neural networks (RNNs) have been used for superior exploration of the temporal information of sequential data. Unlike FFNNs, RNNs permit recurrent connections to create loops, thus allowing a “memory” of previous inputs to remain in the interior state of the network. However, in the training process of the RNN, two major problems exist, namely vanishing and exploding gradients. To tackle these issues, long short-term memory (LSTM) models have been utilized [12]. The LSTM is a modified RNN that was designed to store and access information in a sequence of long periods. The feature maps in the convolution layer are connected to multiple contiguous frames in the preceding layer by this construction, which allows capturing motion information [13].

The OAM mode recognition is an essential and very important task in OAM-based communication systems [14], [15]. The recognition technique must have the ability to mitigate the effect of turbulence on OAM modes. A high-quality recognition process requires a large training set and a more complex network structure in the classification model. The efficient OAM mode recognition improves the OAM-SK demodulation. With successful OAM mode recognition and demodulation, the capacity of the photonic system is enhance.

Due to the orthogonality between different OAM modes, the vortex beam carrying OAM modes has attracted a great attention in the optical communication field. To recognize OAM modes, several popular methods were used to demodulate incident LG beams into Gaussian-like beams, including grating, spiral phase plates (SPPs), and computer-generated holograms (CGHs) equipped with spatial light modulators (SLMs). Incident beams bearing a variety of OAM states can be demodulated into Gaussian-like beams, and the initial information can be retrieved by looking for a bright spot on the camera screen [14], [15]. In addition, a method based on coordinate transformation was also proposed for separating and recognizing OAM modes [15]. However, an OAM-based FSO communication system is subject to atmospheric turbulence (AT) effect, which may unexpectedly distort the phase front of a light beam, cause channel crosstalk, and deteriorate the communication system performance. Several turbulence mitigation techniques, such as multi-input multi-output (MIMO) equalization, adaptive optics (AOs), and phase correction algorithms, such as the Gerchberg-Saxton (GS) algorithm, have been investigated to improve the accuracy of the transmitted beams and reduce the turbulence-induced distortion [14], [15]. To further improve the performance of OAM modulation detection, efficient artificial neural network (ANN), CNN, RNN, and K-nearest neighbor (KNN) based recognition methods have been proposed. With these methods, intensity images of received beams are directly

recognized and the corresponding OAM mode information is successfully obtained [15]–[18].

To the best of our knowledge, it is the first time in this paper to use coding with convolution codes, Reed Solomon (RS) codes, Turbo codes, Bose-Chaudhuri-Hocquenghem (BCH) codes, or Low-Density Parity Check (LDPC) codes on 3D chaotic interleaved VFs over an N -OAM-SK-FSO system, where N refers to the number of employed OAM states. In addition, we deduce the optimum operating parameters in terms of coding type, N -OAM states and VFs. The 3D-CNN and the convolution recurrent neural network (CRNN) techniques with and without GPU are utilized to further enhance and accelerate the detection process and to study the classification and prediction performance under different conditions.

The major motivation of this work is studying the impact of 2D and 3D chaotic interleaving (CI) of coded VFs on the performance of the N -OAM-SK-FSO system. The performance of both deep learning (DL) techniques, namely 3D-CNN and CRNN is assessed by computing the mini-batch and validation accuracy, the training and validation loss, and the processing time with and without GPU. We adopt a trial-and-error method in both DL techniques, and the employed datasets are obtained by applying different scenarios, such as changing code types, using different N -OAM states, applying multiple classes, and employing various VFs. The performance metrics are estimated on these different datasets until reaching the optimum ones for all scenarios. Finally, a comparison is presented between the two proposed DL techniques and the other traditional techniques.

The main contributions of this work are as follows:

- Offering a 3D chaotic interleaving technique for enhancing the performance of coded VF transmission through N -OAM-SK-FSO model.
- Introducing two alternative effectual DL techniques, namely 3D-CNN and CRNN, for enriching the OAM-SK detection efficiency and enhancing the classification performance of the model for different coded and encrypted VFs.
- Studying the different metrics of deep neural networks with different datasets until getting the optimal dataset and the best coded encrypted VFs.
- Monitoring of the bias values by comparing validation and training losses for various dataset cases to get the best ones without overfitting or underfitting.
- Implementation of parallel processing with GPU to accelerate the interleaving process and determine the optimum DL technique and parameters.

The structure of this paper is as follows. In Section II, we present the related work about OAM-SK transmission through the AT channel. In Section III, we describe, in detail, the propagation of the different 3D chaotic interleaved VFs through N -OAM-SK-FSO system. In Section V, we explain two adaptive demodulation mechanisms for OAM-SK using two DL techniques, namely CRNN and 3DCNN, and then we use them for OAM mode detection. Finally, Section VI is devoted to the conclusion.

II. RELATED WORK

In [19], a progressive Turbo encoder/decoder construction and an OAM mapping scheme have been implemented to reduce the CNN misclassification and to efficiently reduce the bit error rate (BER) under different AT conditions. The mitigation of AT effects has been studied theoretically using phase patterns, and the experimental results show that the inter-channel crosstalk is diminished by up to 21 dB as shown in [20]. In [21], an AT correction approach based on DL has been proposed to correct and increase the mode purity of the distorted vortex beam from 39.52% to 98.34% under strong AT. It has been shown that OAM-carrying beams produced by impressing only a helical phase structure on a simple Gaussian beam or truncated plane-wave beam are not optimal for AT-impaired practical OAM-based FSO communication systems. In addition, the control of the initial beam-field amplitude is useful for weakening of the OAM mode scrambling of [22]. In our previous work [23], an alternative effectual CNN architecture designed on the basis of trial and error till getting the optimum network parameter and hyperparameter values was introduced. It yields the highest accuracy and mean average precision (MAP), and the largest area under the curve (AUC) value for the different optimizers. The goal is to get the optimum model using 2D CI with Turbo and LDPC coded images through a 16 OAM-SK system.

In [24], an OAM multiplexing method based on the combination of the Mach-Zehnder interferometer (MZI) implanted pair of dove prisms (DP) (MZI-DP) and a complex phase grating (CPG), was used to sort and detect multiple OAM modes, simultaneously. The OAM modes can be sorted into odd and even topological charge values, and analysis of several OAM modes can be performed without overlapping between adjacent modes. In addition, 50 different OAM modes can be successfully demultiplexed, while controlling their propagation paths, demonstrating high recognition rates and superior efficiency. In [25], a high-density OAM analyzer made up of two OAM mode converters and a modified MZI was presented. The MZI is made up of a standardized MZI and a pair of DPs that effectively avoids the inherent overlap of two adjacent OAM modes in the OAM mode converter. As a result, the scheme precisely avoids the overlapping effect, while also solving the problem of minimal discrimination resolution, allowing for the study of high-density OAM modes. A comparison between the efficiency of different multiplexing methods (coherent superposition or incoherent superposition) and de-multiplexing methods (2D-Dammann fork grating (DFG) or multiplexing phase hologram (MPH)) for vortex beams was presented in [26]. To recover information, the detector uses the optical intensity at a specific location to decide if a data bit is '0' or '1'. So, an image can be delivered and retrieved pixel by pixel with less errors by this way.

However, in this work, compared to our previous work [23], multiple error detection and correction codes with 3D CI technique are used to improve the efficiency of VF communication via the N -OAM-SK-FSO system.

Two effective DL techniques, namely CRNN and 3DCNN are used as adaptive demodulators and a study of the different performance metrics of them is presented on multiple datasets. The aim is to determine the optimum implemented code, OAM state, number of classes, and the supreme DL technique with and without GPU. This is performed to ensure the highest classification and prediction accuracy with the least runtime, and the minimum training and validation losses to mitigate the bias effect, and minimize the BER.

III. MULTIPLE VF TRANSMISSION THROUGH THE PROPOSED SYSTEM

Figure 1 demonstrates the implementation of the proposed system for VF communication. In the proposed system, intensity modulation and direct detection (IM/DD) are employed, where the transmitter modulates the optical power and the receiver consists of a single photodiode, a clock recovery module, and a symbol decision unit. At the transmitter, the encrypted VFs are passed through one of five coding selections (Turbo, LDPC, RS, Convolution, and BCH). Next, the modulation process is performed through an N-OAM-SK mapping scheme according to the value of N as shown in Fig. 2 thanks to OAM beam orthogonality. Each OAM state is represented by a topological charge and a binary value as shown in Table 1. After propagation through the turbulence channel, the blurred images are passed through a charge coupled device (CCD) camera [19]. After that, the received images are passed through a switch with two options. The first option is the utilization of one of two DL-based adaptive demodulation techniques, namely 3DCNN, and CRNN. The second option is obtaining the electrical amplitude signals from the optical brightness of the CCD image sensor. At every time, only a single beam reaches the receiver, i.e., the system is single-input single-output (SISO).

The demapping process is performed in consistence with the conjugate mode sorting technique to clarify the mode numbering of the detected beam [27]. After demapping, the decoding process is performed according to the previously

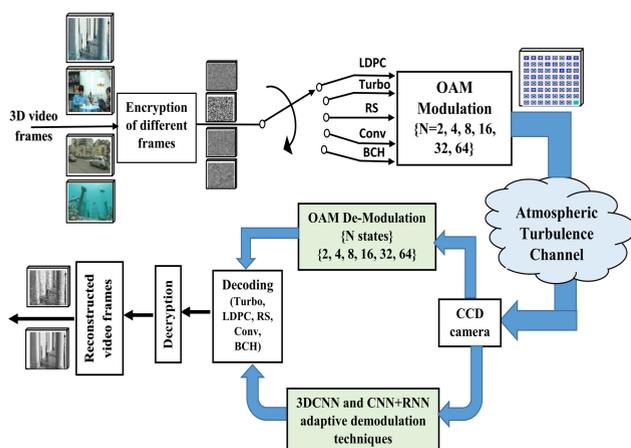


FIGURE 1. Video frame communication through the proposed system.

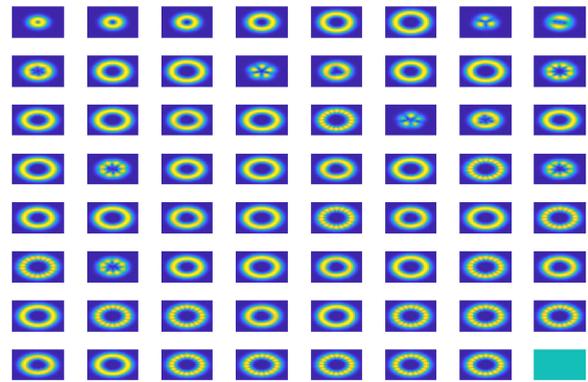


FIGURE 2. Different states of OAM-SK for $N = 64$.

selected code. Finally, a 3D chaotic deinterleaver is applied on the decoded data to decrypt it and restore the original 3D VF to constitute various videos. Thermal noise is caused by the load resistor and amplifier noise of the photodetector. It typically dominates shot noise in mild conditions, where high-sensitivity receivers are not needed. So, the used link is thermally limited [28]

The response speed of the SLM and the CCD limits the OAM-SK image recognition (IR) capability. Long data processing times have a significant impact on performance due to the slow scanning speed and frame rate of CCD-based cameras. Despite these limitations, SLM and CCD cameras outperform IR-based OAM systems in several applications. They are used to achieve ultra-high capacity and encryption security in the OAM holography system [29]. The CCD camera is integrated into a robust OAM mode fiber laser device with an all-polarization-maintaining fiber (PMF) structure, resulting in an OAM mode fiber laser that is resistant to environmental disruptions and does not require the use of a polarization controller [30]. In a recently constructed annular core photonic crystal fiber (AC-PCF), the CCD cameras were used in stable OAM propagation and cylindrical vector beams. Several applications in space-division multiplexing, optical sensing, and tracking are possible with this method [31]. These cameras were used in OAM-SK underwater wireless optical communication (UWOC) systems in oceanic turbulence channels [32]. The wave front distortion of optical beams induced by atmospheric turbulence is measured using CCD cameras in low-frequency heterodyne interferometry [33].

A. THE ENCRYPTION PROCESS OF 3D CHAOTIC INTERLEAVED CODED VFs

To encode the different 3D videos, the 3D multi-view video coding prediction structure (3D-MVC PS) is employed due to its efficient coding and decoding performance as shown in Fig. 3. It incorporates eight separate time pictures in a group of pictures (GOPs). The horizontal axis represents the temporal axis, while the vertical axis represents numerous camera views. For example, depending on the motion compensation prediction (MCP), the S_0 view is only coded using time correlation [34]. The MCP is also used to code the other

TABLE 1. The representation of the different 64 OAM states.

States	Topological charge	Binary sequence	States	Topological charge	Binary sequence
1	{1}	100000	33	{-2,3,-8}	011010
2	{-2}	010000	34	{-2,3,10}	011001
3	{3}	001000	35	{-2,-5,-8}	010110
4	{-5}	000100	36	{-2,-5,10}	010101
5	{-8}	000010	37	{-2,-8,10}	010011
6	{10}	000001	38	{-2,3,-5}	001110
7	{1,-2}	110000	39	{3,-5,10}	001101
8	{1,3}	101000	40	{3,-8,10}	001011
9	{1,-5}	100100	41	{-5,-8,10}	000111
10	{1,-8}	100010	42	{1,-5,10}	100101
11	{1,10}	100001	43	{1,-2,3,-5}	111100
12	{-2,3}	011000	44	{1,-2,3,-8}	111010
13	{-2,5}	010100	45	{1,-2,3,10}	111001
14	{-2,-8}	010010	46	{1,-2,-5,-8}	110110
15	{-2,10}	010001	47	{1,-2,-5,10}	110101
16	{3,-5}	001100	48	{-2,3,-5,-8}	011110
17	{3,-8}	001010	49	{-2,3,-5,10}	011101
18	{3,10}	001001	50	{3,-5,-8,10}	001111
19	{-5,-8}	000110	51	{1,-2,-8,10}	110011
20	{-5,-10}	000101	52	{1,3,-5,10}	101101
21	{-8,-10}	000011	53	{1,3,-8,10}	101011
22	{1,-2,3}	111000	54	{-2,-5,-8,10}	010111
23	{1,-2,-5}	110100	55	{1,-2,3,-5,-8}	111110
24	{1,-2,-8}	110010	56	{1,-2,3,-5,10}	111101
25	{1,-2,-10}	110001	57	{1,-2,3,-8,10}	111011
26	{1,3,-5}	101100	58	{1,3,-5,-8,10}	101111
27	{1,3,-8}	101011	59	{1,-2,-5,-8,10}	110111
28	{1,3,-10}	101001	60	{-2,3,-5,-8,10}	011111
29	{1,-5,-8}	100110	61	{1,-5,-8,10}	100111
30	{1,-5,-10}	100101	62	{1,3,-5,-8}	101110
31	{1,-8,-10}	100011	63	{1,-2,3,-5,-8,10}	111111
32	{-2,3,-5}	011100	64	Zeros	000000

even views as $S_2, S_4,$ and S_6 . The main key frames, on the other hand, are coded using inter-view prediction (disparity compensation prediction (DCP)). Both inter-view and temporal estimations (DCP + MCP) are used simultaneously in the odd views $S_1, S_3,$ and S_5 to improve the coding efficiency. The 3DV view is studied from the perspective of the 3DV frame basic locality. As shown in this figure, the odd ($S_1, S_3,$ and S_5) views are referred to as E views, the even ($S_2, S_4,$ and S_6) views as O views, and the S_0 view as I view. Based on the suggested 3D-MVC PS GOPs, the final view could be even or odd. It is suggested in this paper to be an O view [34], [35].

The CI technique is selected based on the weak correlation between output samples, and the elimination of either huge memory or wide bandwidth. A major criterion of it is the interleaver depth that is defined as the minimum separation between any pair of symbols, whether adjacent or separated with a certain distance at the output [36]. When the length of a burst of errors is less than the interleaver depth, after de-interleaving, the two symbols affected by the burst cannot be adjacent. So, a simple isolated error correction code can be used to enhance the performance. In the presented 3-D CI, the goal is to increase the randomness, security, depth, and reliability of the N -OAM-SK model. The 3D-MVC coding scheme is implemented to compress the transmitted 3DV sequences due to its efficient coding and decoding. Several simulation tests on the standard well-known 3DV test sequences, e.g., Shark, PoznanStreet,

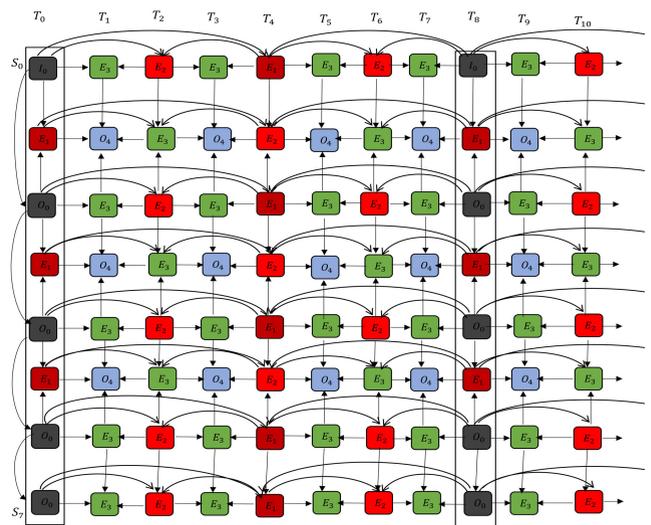


FIGURE 3. Video coding and prediction structure.

Dancer, and Newspaper, have been carried out with a resolution of 1920×1088 [36]. The utilized 3D video sequences have different spatial and temporal characteristics. The Dancer and PoznanStreet videos are fast-moving sequences. The Shark stream is a moderate animated video. The Newspaper sequence is an intermediate slow motion video. For each 3DV sequence, eight views with 250 frames in each view are coded and the frame rate is 20 frames/sec. For each sequence,

the coded 3D H.265/HEVC bit streams are produced by employing the reference HEVC test Model (HM) codec [35].

To generate the data, the first step is the segmentation process of different 3DVs to get the different 3DV frames. The different VFs are then encrypted using 3D chaotic map [36] according to:

$$\begin{aligned} E_{j+1} &= \alpha \cdot T_j^3 + \beta \cdot P_j^2 \cdot E_j + \gamma \cdot E_j(1 - E_j) \\ P_{j+1} &= \alpha \cdot E_j^3 + \beta \cdot T_j^2 \cdot P_j + \gamma \cdot P_j(1 - P_j) \\ T_{j+1} &= \alpha \cdot P_j^2 + \beta \cdot E_j^2 \cdot T_j + \gamma \cdot T_j(1 - T_j) \end{aligned} \quad (1)$$

where E , P , and T are the original locations of image pixels, $\alpha = 0.0125$, $\beta = 0.0157$, $\gamma = 3.7700$, and $j \in \{1, 2, \dots, 80000\}$. The initial conditions are $E_1 = 0.2350$, $P_1 = 0.3500$, and $T_1 = 0.7350$. The histogram equalization is then used to prepare the image for use, and this is done according to:

$$\begin{aligned} E &= \text{integer}(E \times \Upsilon_2) \pmod{\Upsilon} \\ P &= \text{integer}(P \times \Upsilon_4) \pmod{\Psi} \\ T &= \text{integer}(T \times \Upsilon_6) \pmod{256} \end{aligned} \quad (2)$$

where Υ is the number of rows, Ψ is the number of columns, and Υ_2 , Υ_4 , and Υ_6 are random numbers larger than 10,000 [36]. Then, a column-row rotation is accomplished followed by an XOR on the subsequent shifted image and chaos operation to develop the encrypted image. Finally, the coding process is applied on the encrypted image according to the selected switch as shown in Fig. 1. After that, a type of error-correction and detection codes such as Turbo, LDPC, and RS is used.

B. CODING PROCESS

Coding is one of the fundamental techniques that makes near-capacity operation conceivable. Widespread possibilities for error detection and correction can be realized by encoding and decoding the transmitted data. Some types of error-correction codes such as convolution codes, RS codes, Turbo codes, BCH codes, and LDPC codes are used [37].

C. TRANSMISSION OF DIFFERENT CODED VFs THROUGH N-OAM-SK OVER AT CHANNEL

After coding of the interleaved VFs, the mapping process is performed by exchanging every $\log_2 N$ bits with one of N OAM states as reported by the different states. The different OAM states are generated according to the following equation [19], [38]:

$$\begin{aligned} \kappa_{LG(l_n, p)}(r, \theta, z)_n &= \frac{D}{\left(1 + \frac{z^2}{(\pi\omega_0^2/\lambda)^2}\right)^{1/2}} e^{-\frac{ikr^2z}{2(z^2 + (\pi\omega_0^2/\lambda)^2)}} \\ &\cdot e^{-\frac{r^2}{\omega^2(z)}} \left(\frac{r\sqrt{2}}{\omega(z)}\right)^{|l|} l_p^{|l|} \left(\frac{2r^2}{\omega^2(z)}\right) e^{-il\theta} e^{ikz} \\ &\cdot e^{i(2p+|l|+1)\tan^{-1}\left(\frac{z}{\pi\omega_0^2/\lambda}\right)} \end{aligned} \quad (3)$$

where D is a normalization constant, r is the radial distance from z , l is the intertwined helical phase front, p is the radial index, $p + 1$ is the number of circular or ring regions, (r, θ, z) are the cylindrical coordinates, $k = 2\pi/\lambda$ is the wave number that refers to the number of radians per unit distance, and it is sometimes called ‘‘angular wave number’’, λ is the optical wavelength, and $l_p^{|l|}$ is the generalized Laguerre polynomial. The beam radius of the fundamental Gaussian beam at distance z is given by [19]:

$$\omega(z) = \omega_0 \sqrt{1 + (z/z_R)^2} \quad (4)$$

where ω_0 is the beam waist at $z = 0$.

In the simulation process, the beam waist ω_0 at the transmitter for all beams is set to 1.5 cm to ensure a minimum beam waist at the receiver plane [38]. The propagation distance is from 200 to 1000 m and the numbers of OAM channels are 2, 4, 8, 16, 32, and 64. Also, the wavelength λ is equal to 1550 nm. On the other hand, the optical receiver is assumed to be large enough to collect all obtained OAM beams.

The generation process of the OAM modes with state $l \in \{1, -2, 3, -5, 8, 10\}$ in practical implementation can be accomplished using cylindrical lenses, SPP, SLM, holographic plate, liquid-crystal devices, metamaterials, and CGH [39], [40]. In communication experiments involving OAM beams at various wavelength ranges, SLMs are the most widely used instruments [38]. The obtained OAM-SK modulated signals are transmitted over the AT channel. The transmitted OAM beams are exposed to atmospheric turbulence (AT) in real life communication scenarios, which causes spatial variations in the air refractive index. When OAM beams travel through a turbulent atmosphere, they experience phase-front distortions, beam spread, and wandering. Moreover, modal crosstalk occurs, when the power of a signal carried by an OAM mode is spread across other modes as shown in Fig. 4. The distribution of optical power to other OAM modes is low in the presence of weak turbulence. However, in a strong turbulence regime, optical power spreads widely, and there is a significant power leakage to other OAM modes. The latter is mode-dependent, and it causes a loss of orthogonality between OAM modes, resulting in loss disparities known as mode-dependent loss (MDL), which causes system performance degradations [38], [41], [42].

In this work, we model the channel impact as a gamma-gamma turbulence model given by [43], [44]:

$$\psi(h) = \frac{2(\mu\nu)^{\frac{\mu+\nu}{2}}}{\Gamma(\mu)\Gamma(\nu)} h^{\frac{\mu+\nu}{2}-1} K_{\mu-\nu} \left(2\sqrt{\mu\nu}h\right) \quad (5)$$

where μ and ν represent the effective number of large-scale and small-scale eddies of the scattering process, respectively. The values of μ and ν are calculated with the help of the following equations [19], [43]:

$$\mu = \left[\exp \left(0.49\rho_v^2 / (1 + 1.11\rho_v^{12/5})^{7/6} \right) - 1 \right]^{-1} \quad (6)$$

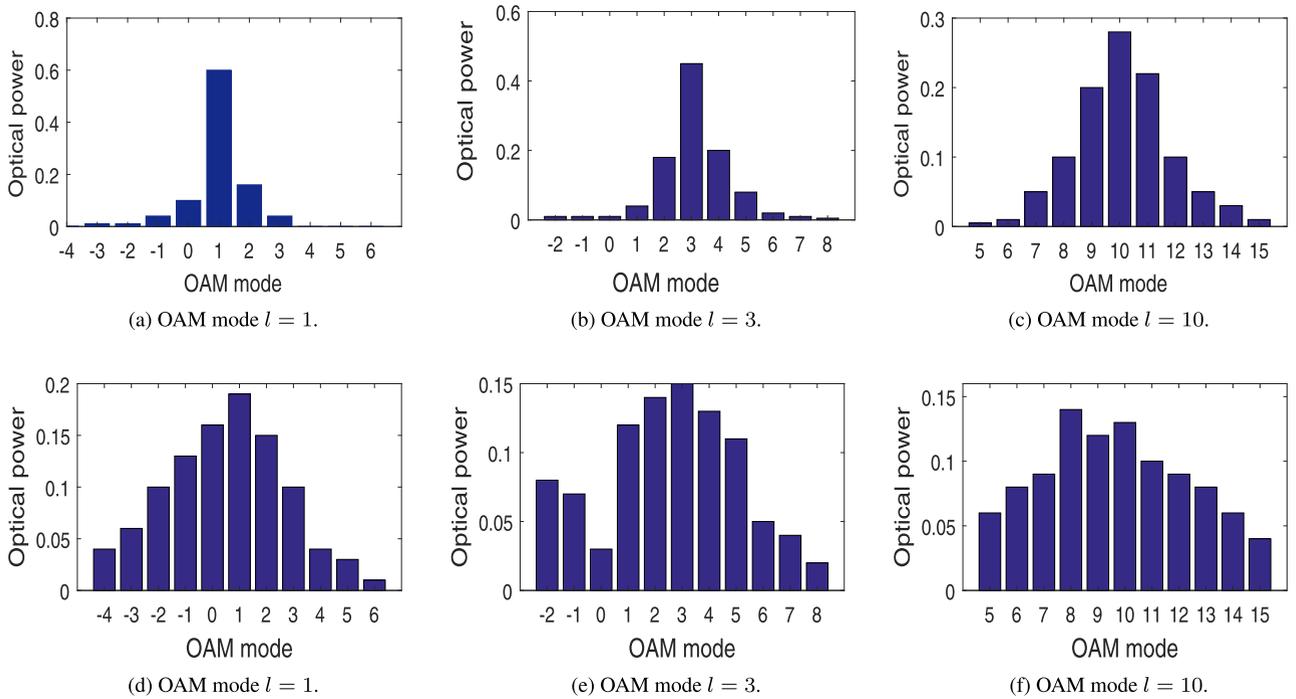


FIGURE 4. The received power distribution for OAM modes $l = 1, 3,$ and 10 for weak and strong atmospheric turbulence regimes: (a):(c) Weak AT, (d):(f) Strong AT.

$$v = \left[\exp \left(0.51 \rho_v^2 / \left(1 + 0.69 \rho_v^{12/5} \right)^{5/6} \right) - 1 \right]^{-1} \quad (7)$$

$$= \begin{cases} \int |\kappa_{LG(l_m,p)}|^2 r dr d\theta; & \text{for } m = n, \\ 0; & \text{for } n = 2, 4, 8, 16, 32, 64. \end{cases} \quad (9)$$

where $\rho_v^2 = 1.23 c_n^2 k_0^7 D_p^{11/6}$ is the variance of the irradiance fluctuations [19], [38], [48], h is the normalized received irradiance, $K_n(\cdot)$ is the modified Bessel function of the 2nd kind of order n , $\Gamma(\cdot)$ represents the Gamma function, c_n^2 is the air refractive index structure constant, used to describe the strength of the AT, and D_p is the propagation distance.

Due to the presence of the modified Bessel function in (5), a significant mathematical complexity in directly dealing with the $\psi(h)$ of the channel exists. Therefore, we use Meijer G-function to represent the modified Bessel function, and $\psi(h)$ becomes [1], [43]:

$$\psi(h) = \frac{2 \cdot (\mu\nu)^{\frac{\mu+\nu}{2}} h^{\frac{\mu+\nu}{2}-1} G_{0,2}^{2,0} \left(\mu\nu h \left| \begin{matrix} \cdot, \cdot \\ \frac{\mu-\nu}{2}, \frac{\nu-\mu}{2} \end{matrix} \right. \right)}{\Gamma(\mu)\Gamma(\nu)}. \quad (8)$$

Then, after the turbulence channel, the reverse operation of the transmitter can be carried out at the receiver side using the same device to convert the incoming OAM mode back to a Gaussian beam. The idea is to apply an optical scalar product measurement on the image plane of a Fourier transforming lens between the incident OAM beam and a CGH with the conjugate phase [38], [43].

$$\begin{aligned} \widehat{\kappa}(r, \theta, z)_m &= \langle \kappa_{LG(l_m,p)}(r, \theta, z)_m \kappa_{LG(l_n,p)}^*(r, \theta, z)_n \rangle \\ &= \int \kappa_{LG(l_m,p)}(r, \theta, z)_m \kappa_{LG(l_n,p)}^*(r, \theta, z)_n r dr d\theta \end{aligned}$$

where l_n is the topological charge of the originally-used OAM states, l_m is the topological charge of the received corrupted data. $\kappa_{LG(l_m,p)}(r, \theta, z)_m$ refers to the cylindrical coordinates of the m^{th} received corrupted beam.

The result of this product is estimated according to one of the two subsequent decisions [27]:

- 1) If the real and the imaginary terms equal zero, then the current mode number is 2, 8, 4, 16, 32, or 64 according to the value of N .
- 2) If the imaginary term only equals zero at any time, then it is the state of interest, and this is the correct mode number.
- 3) If the real and imaginary parts exist and are not equal to zero, then it is the case of $m \neq n$ and n not equal to any index of the applied states.

The detection process is performed using the conjugate light field detection method. The estimation of l is performed based on the orthogonality property of OAM states. Assuming that the LG beam after turbulence channel is $\kappa_{LG(l_m,p)}(r, \theta, z)_m$, we can compute the product between the selected LG beam and the conjugate of each of the used modes $\kappa_{LG(l_n,p)}^*(r, \theta, z)_n$, for beam recovery, where m and n characterize the m^{th} and n^{th} beams [27]:

After demapping all received data to the different OAM states, the decoding process is performed and followed by the decryption process of the recovered data to yield the original bit stream, and then get the original videos. Now, the absolute

error rate is obtained by multiplying the computed error rate of the introduced model with the average error rate of the turbulence channel. The BER value is calculated by taking the average value:

$$BER = \widehat{\kappa}(r, \theta, z)_m - \kappa_{LG(l_n, p)}(r, \theta, z)_n \quad (10)$$

Finally, we calculate the image quality metrics between the reconstructed image and the original image as follows [37], [44]:

- MSE (Mean Square Error):

$$MSE = \frac{1}{\Upsilon\Psi} \sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (x_{i,k} - y_{i,k})^2 \quad (11)$$

- PSNR:

$$PSNR = 10 \log \frac{(2^w - 1)^2}{MSE} \quad (12)$$

- NX (Normalized Cross Correlation):

$$NX = \frac{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (x_{i,k} \times y_{i,k})^2}{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (x_{i,k})^2} \quad (13)$$

- AD (Average Difference):

$$AD = \frac{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (x_{i,k} - y_{i,k})^2}{\Upsilon\Psi} \quad (14)$$

- MD (Maximum Difference):

$$MD = \max(|x_{i,k} - y_{i,k}|) \quad (15)$$

- SC (Structural Count):

$$SC = \frac{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (x_{i,k})^2}{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} (y_{i,k})^2} \quad (16)$$

- NAE (Normalized Absolute Error):

$$NAE = \frac{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} |x_{i,k} - y_{i,k}|}{\sum_{i=1}^{\Upsilon} \sum_{k=1}^{\Psi} |x_{i,k}|} \quad (17)$$

Here, Υ and Ψ are the dimensions of the monochromatic image, x is the original image, y is the noisy image, and w is number of bits per sample.

IV. SIMULATION RESULTS OF N-OAM-SK-FSO MODEL

In the simulation experiments, the value of the refractive index structure function c_n^2 is set according to the different turbulence strength cases: low, moderate, and strong with values of 10^{-15} , 10^{-14} , and 10^{-13} , respectively [38]. Also, the SNR values are changed from 0 to 20 dB and the power of the signal is set to a fixed value relevant to the transmitted data amplitude. The noise power is altered according to the SNR. The different encrypted coded VFs are used through OAM-SK-FSO model at different values of turbulence strength, propagation distance, and SNR. In addition, the performance metrics of image quality are measured for different 3D encrypted coded VFs using 16-OAM-SK. The simulation results indicate the following:

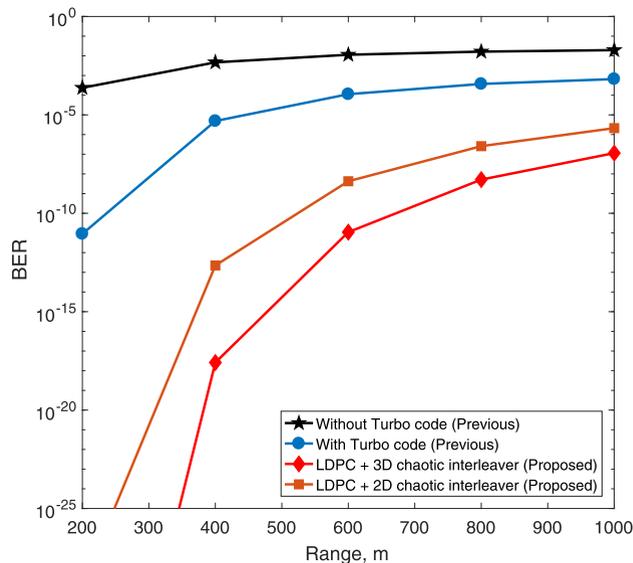


FIGURE 5. Comparison between the previous and proposed models for different codes.

A comparison between the proposed work and other previously known research works is indicated as follows:

- 1) Figure 5 demonstrates a comparison between the previous model in [19] and the proposed model with different values of propagation distance. It is clear from the figure that the proposed model with LDPC coding has a lower BER compared to those of the previous model by a factor of about 10^{-4} after using the 2D CI and 10^{-5} after using the 3D CI.
- 2) Figure 6 indicates a BER comparison between the previous model in [45] and the proposed model using LDPC codes for different SNR values. It is clear from the figure that the proposed model at 10 dB has a lower BER compared to that of the previous model by 7, and 10 orders of magnitude using 2D, and 3D CI, respectively.
- 3) Table 2 indicates the efficiency of the proposed model over other models. From Table 2, it is clear that the proposed model using 3D CI is better at an SNR value of 7 dB by 3, 4, 6, and 6 orders of magnitude compared to the models in [46]–[49], respectively.

It is shown from Figs. 7–9 that the BER value increases with the increase of propagation distance, and turbulence strength but decreases with the increase of the SNR for the 3D chaotic interleaved coded VFs. In this figure, we depict the BER performance for the 16 OAM-SK-FSO system using coded chaotic interleaved VFs for different values of

TABLE 2. Comparison between the BERs of the proposed and previous models.

References	[46]	[47]	[48]	[49]	Proposed model
BER	10^{-9}	10^{-8}	10^{-6}	10^{-6}	10^{-12}

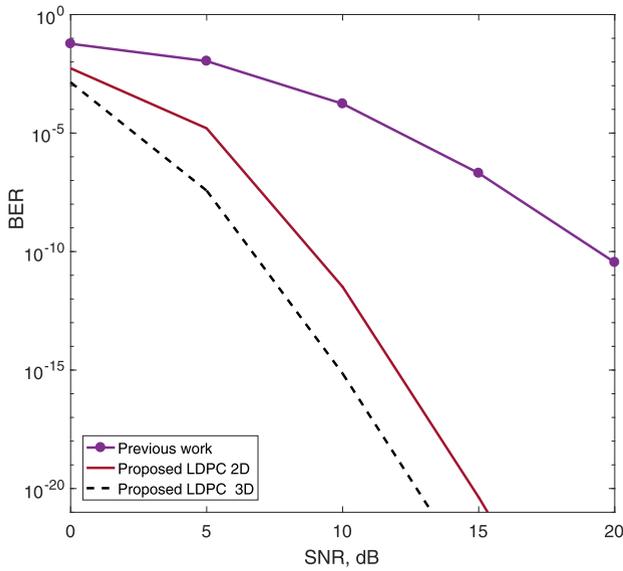
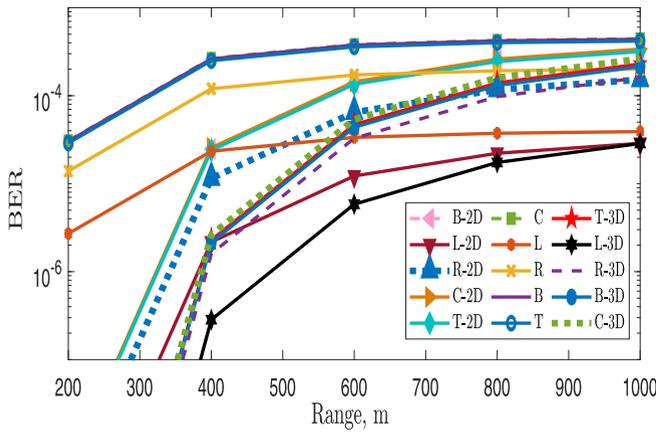


FIGURE 6. Comparison between the previous and the proposed models using LDPC code.

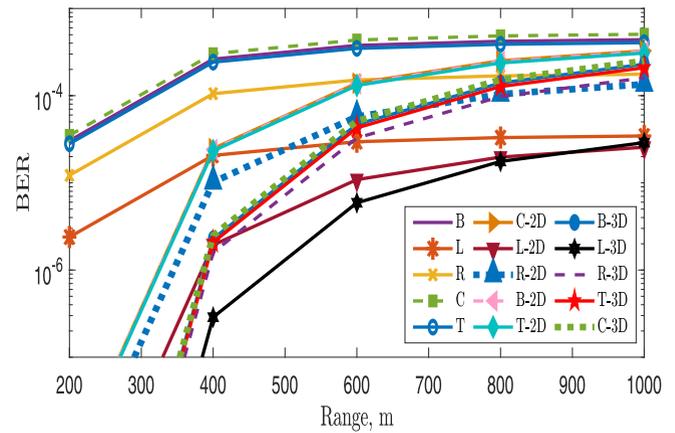
propagation distance, turbulence strength, and SNR. For all cases, LDPC-coded Newspaper VFs have the minimum BER

compared to other interleaved coded VFs. The BCH and convolution codes achieve the maximum BER for all used VFs. Due to space limitations, only the case of 16-ary OAM-SK is included as shown in these figures as it has the best performance in comparison with 32 and 64 states due to the existence of several repeated states as shown in all obtained states for the different values of N . Fig. 2 shows a state of them at $N = 64$.

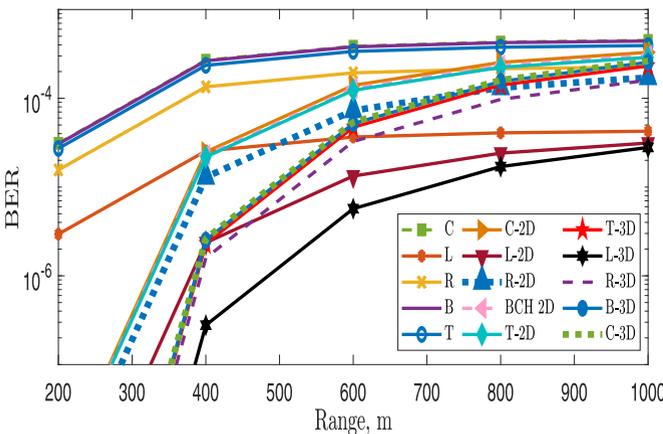
In Table 3, we introduce a comparison between the different metrics of image quality for coded VFs and different states (L:low, M:medium, S:strong) of turbulence strength. The usage of 1, 2, and 3 in the table refers to without interleaver, 2D, and 3D CI cases. In all turbulence states, LDPC code is better than RS, Turbo, Convolution, and BCH codes by almost 1, 1.5, 2.5, and 3 respectively. Using 3D CI at low turbulence strength enhances the model performance by nearly 3 dB. Through medium turbulence strength, the existence of 2D and 3D CI improves the OAM-SK model by approximately 2 and 3 dB, respectively. At strong turbulence strength, the 3D CI improves the model performance by nearly 5 dB. From the table, we notice that the LDPC code is the optimum code



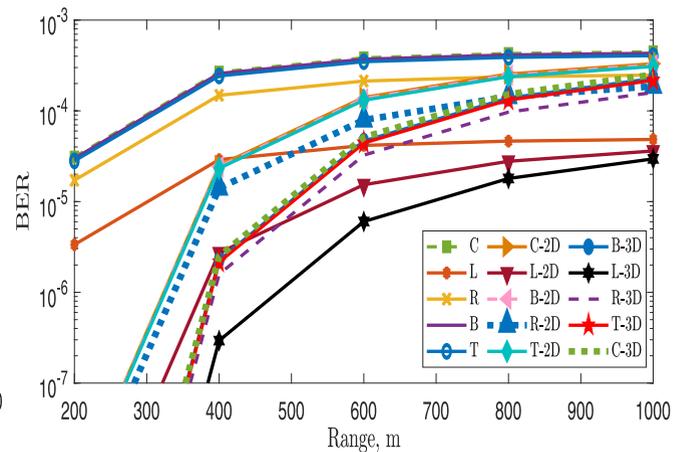
(a) Shark.



(b) Dancer.



(c) Newspaper.



(d) Poznanstreet.

FIGURE 7. BER using 4 video frames at 16 OAM-SK for different propagation distance values. (T: Turbo code, L: LDPC code, R: RS code, B: BCH code, and C: Convolution code).

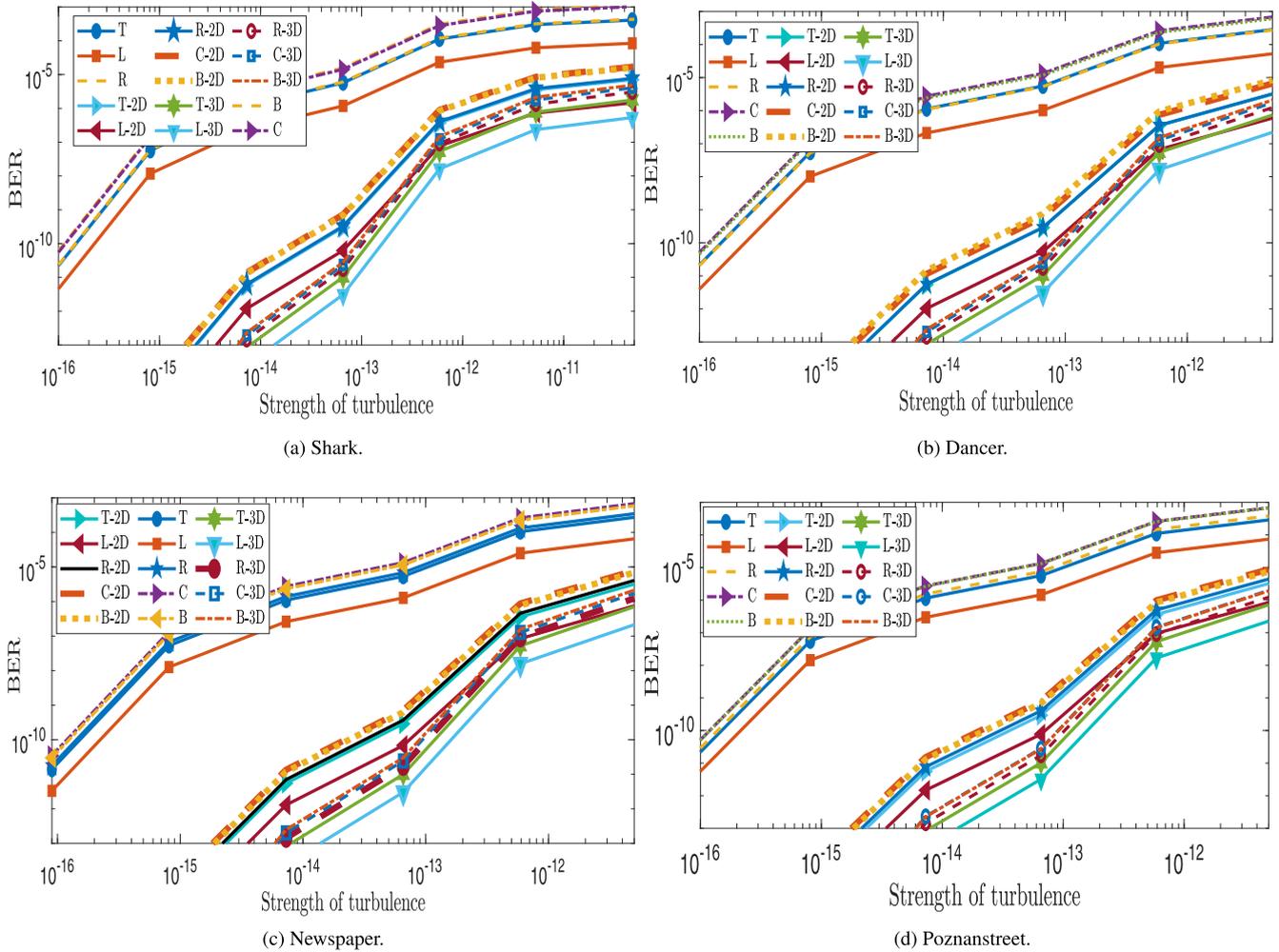


FIGURE 8. BER using 4 video frames at 16 OAM-SK for different turbulence strength values. (T: Turbo code, L: LDPC code, R: RS code, B: BCH code, and C: Convolution code).

that achieves the best image quality for different turbulence strengths followed using RS, Turbo, Convolution, and finally BCH code.

In Table 4, we make a comparison between the different metrics of image quality using coded OAM-SK and different values of propagation distance. The existence of 1, 2, and 3 in the table refers to the cases of without interleaver, 2D, and 3D CI. Through different values of propagation distance, LDPC code is better than RS, Turbo, Convolution, and BCH code by almost 1, 1.5, 2.5, and 3, respectively. Using CI of 2D or 3D nature enhances the model performance by nearly 3 dB. The LDPC code has the optimum values of PSNR, MSE, NK, NAE, SC, MD, and AD for all different values of propagation distance, while BCH code has the lowermost quality metrics compared to those of other codes.

After that, a comparison is introduced between the different image quality metrics under different propagation distances and turbulence strengths. From Tables 3 and 4, it is shown that:

- 1) LDPC code is the optimum used code with the highest PSNR and the lowest MSE for all turbulence strengths and propagation distances with and without CI.
- 2) LDPC code has the least values of MD, SC, NAE, NX and AD, which means that it achieves only small error values and the reconstructed and the reference video frames are nearly the same, which means high similarity.
- 3) The RS and Turbo codes are very close to LDPC code in performance, but they do not reach the same level of its performance.
- 4) The BCH and convolution codes achieve the least image quality metrics for all turbulence strengths and propagation distances.

The explanations of the above points are as follows [37]:

- LDPC codes have superior error-correction capabilities and low error floor. They also benefit from the simple implementation of encoders/decoders. They per-

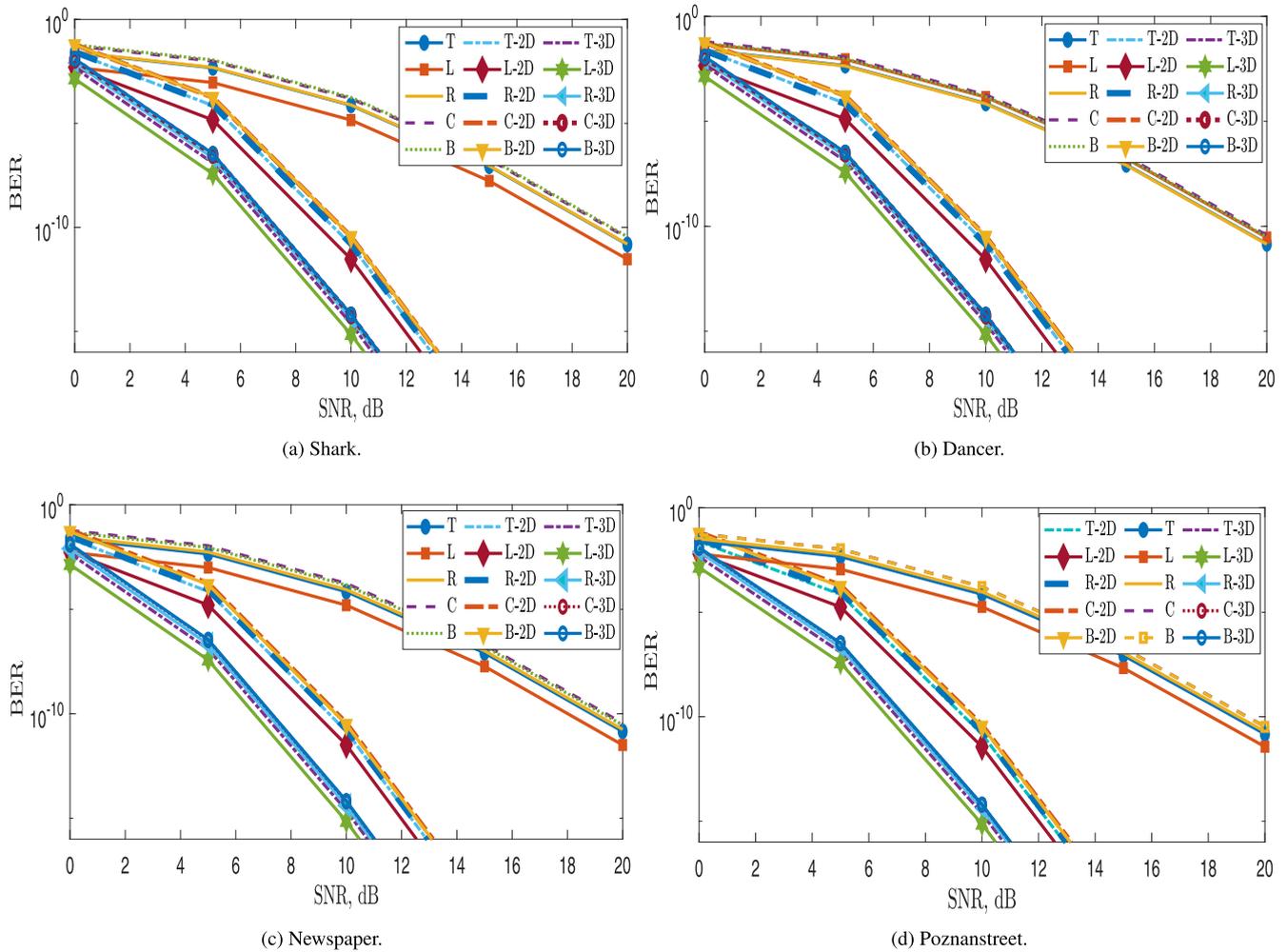


FIGURE 9. BER using 4 video frames at 16 OAM-SK for different SNR values. (T: Turbo code, L: LDPC code, R: RS code, B: BCH code, and C: Convolution code).

form much better in correcting random errors than other codes.

- For Turbo codes, despite their excellent decoding efficiency, they are hard to implement for parallel decoding. They suffer from relatively high decoding latency and complexity of decoding as they follow interlacing structures and iterative algorithms of decoding.
- RS codes are not good for correcting random errors compared to Turbo and LDPC codes, while RS codes do well with non-binary algebraic structures to correct burst errors. Using RS codes with more parity check symbols typically leads to better burst-error-correcting efficiency, but with more complexity.
- The efficiency of convolution codes is less than those of Turbo and LDPC codes in terms of correction of burst errors. In addition, they puncture bits to achieve high code speeds, which results in losing some useful information. The encoding process of Turbo codes is simple, whereas the decoding process is complicated.

- In BCH codes, the major characteristic is the precise control over the number of correctable symbol errors during the design process. Another property of them is the ease of the decoding process and the simple design of the decoder using a slight little-power hardware.

To enhance the RS and Turbo codes for the proposed design:

- We use a simple circuit and small datasets to improve the performance of these codes.
- We employ a small number of OAM states as 2, 4, and 8, which means low complexity of the model.
- We increase the number of iterations and the interleaver size.

V. ADAPTIVE DETECTION MECHANISM OF OAM-SK USING TWO DL METHODS

The conventional OAM-SK demodulation techniques are unreliable and highly prone to angular fluctuations incorporating effects of variations in both random beam misalignment and angle-of-arrival (AOA) [50]. In addition to the beam

TABLE 3. Comparison between a variety of image quality metrics using different codes at different turbulence strengths.

Codes		Turbo			LDPC			RS			Convolution			BCH		
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
MSE	L	121	70	39	94	41	22	125	78	27	159	98	68	159	93	75
	M	150	81	63	99	54	22	127	80	29	162	107	72	162	104	88
	S	181	136	114	123	93	23	156	84	57	171	124	88	171	116	99
PSNR	L	27	30	32	28	32	35	27	29	34	26	28	30	26	28	29
	M	26	29	30	28	31	35	27	29	33	26	28	30	26	28	29
	S	26	27	28	27	29	35	26	29	31	26	27.2	29	26	27	28
NK	L	0.88	0.95	0.98	0.9	0.98	1.01	0.85	0.98	0.97	0.72	0.98	0.96	0.74	0.96	0.94
	M	0.79	0.94	0.96	0.9	0.96	1.02	0.86	0.97	0.97	0.68	0.93	0.97	0.72	0.92	0.92
	S	0.65	0.84	0.89	0.85	0.9	1.01	0.77	0.96	0.95	0.68	0.87	.94	0.68	0.88	0.91
NAE	L	0.22	0.11	0.08	0.13	0.07	0.07	0.24	0.22	0.04	0.45	0.39	0.13	0.37	0.24	0.21
	M	0.32	0.14	0.13	0.14	0.08	0.07	0.23	0.21	0.04	0.46	0.4	0.14	0.39	0.27	0.25
	S	0.42	0.26	0.25	0.19	0.13	0.07	0.3	0.2	0.09	0.47	0.41	0.17	0.41	0.29	0.3
MD	L	217	172	181	208	145	111	238	204	141	234	237	193	236	222	236
	M	236	201	215	232	173	106	233	204	146	239	239	194	240	202	236
	S	238	228	245	212	210	99	241	204	205	239	234	210	240	217	236
AD	L	14.2	4.1	1.9	14.5	1.3	-4.5	19.8	-1.1	3.8	30.9	-14.4	3.5	32.3	-2.2	1.6
	M	27.1	6.1	3.7	15.6	4.7	-4.9	19.1	0.05	4.1	35.8	-6.3	3.7	35.1	3.3	1.7
	S	47.7	20.4	11.7	23.3	4.5	-4.6	32.2	3.3	6.9	38.7	5	7.6	41	8.4	1.6
SC	L	1.2	1	1	1.2	1	0.95	1.2	0.96	1	1.5	0.85	1	1.5	0.98	1
	M	1.4	1.1	1	1.2	1	0.95	1.2	0.97	1	1.54	0.92	1	1.5	1	1.1
	S	1.8	1.3	1.1	1.3	1.2	0.95	1.4	1	1.1	1.6	1	1.1	1.7	1.1	1.1

misalignment caused by transmitter vibrations, the receiver vibrations also affect the performance of the FSO link by deteriorating AOA fluctuations. Although the larger field of view (FOV) receiver has a greater ability to reduce the deterioration caused by AOA fluctuations, a limited FOV receiver is desirable to achieve higher data rate transmission, but the system performance is degraded by AOA fluctuations [50]. For OAM-SK direct detection, narrower FOV can be used to minimize the accumulated background noise, but it also decreases the amount of detected signal power, when considering AOA fluctuations. To demodulate OAM-SK, the most popular approach is to use SLM, and SPPS which are complicated and costly physical devices. For this reason, an efficient adaptive demodulator is suggested, based on DL. In addition, an improved interleaved coding technique is exploited, and an advanced mapping scheme is proposed to provide better communication performance under different AT conditions with DL techniques [11]. In this section, we employ adaptive demodulation using two DL techniques to classify and predict distinct videos using different datasets of encrypted coded N -OAM-SK states. Several studies have recently been performed on using DL to perform adaptive demodulation of OAM modes. CNN, as a model of DL, has been successfully implemented in the OAM-FSO communication framework, using a multi-layer representation learning technique with

local relation and shared weights to retrieve and recognize the intrinsic features of the raw input images.

Due to the temporal nature of video frames, the usage of CNN only is not effectual. Hence, we use two other classification and prediction models as shown in Fig. 10 to classify different encrypted coded video frames and get a better classification and prediction performance than that of the traditional CNN model. The first model is CNN + RNN and the second model is 3D CNN as illustrated below.

A. CRNN MODEL

This model contains a group of layers: sequence input layer, sequence folding layer, convolution layer, average pooling layer, sequence unfolding layer, flattening layer, 2 dropout layers, 2 bidirectional LSTM (BiLSTM) layers, a batch normalization layer, a single fully-connected layer, and an output layer. The existence of the sequence input layer is to introduce image sequences and permit the convolutional operations to be applied on each VF, autonomously [12]. In the convolutional layer, a group of filter collections are trained and the convolution operation of these filters with the input VFs with size 12×12 gives several convolution outputs, based on the size and number of the kernels, stride, and padding. The convolutional layer is implemented using 10 kernels each

TABLE 4. Comparison between a variety of image quality measures for different codes using different values of propagation distance.

Codes		Turbo			LDPC			RS			Convolution			BCH		
		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
MSE	200	149	69	27	90	56	19	118	78	28	136	103	84	161	95	67
	600	150	78	52	94	60	21	121	79	28	149	112	93	164	97	73
	1000	153	94	81	96	62	21	124	80	30	155	118	97	171	107	76
PSNR	200	26	30	34	29	31	35	27	29	34	27	28	29	26	28	29
	600	26	29	31	28	30	35	27	29	34	26	28	28	26	28	29
	1000	26	28	29	28	30	35	27	29	33	26	27	28	26	28	29
NK	200	0.79	0.96	0.99	0.91	0.97	1.02	0.87	0.97	0.97	0.8	0.96	0.93	0.73	0.95	0.97
	600	0.79	0.95	0.97	0.9	0.96	1.02	0.87	0.97	0.97	0.77	0.91	0.91	0.71	0.95	0.96
	1000	0.78	0.92	0.94	0.91	0.95	1	0.86	0.97	0.97	0.75	0.89	0.91	0.68	0.92	0.96
NAE	200	0.32	0.11	0.05	0.12	0.08	0.06	0.23	0.21	0.04	0.44	0.39	0.24	0.38	0.26	0.13
	600	0.31	0.12	0.1	0.14	0.09	0.07	0.23	0.21	0.04	0.44	0.4	0.27	0.39	0.25	0.14
	1000	0.32	0.14	0.16	0.13	0.09	0.07	0.23	0.22	0.04	0.42	0.41	0.29	0.42	0.27	0.14
MD	200	230	176	180	214	169	100	217	202	144	234	233	236	239	208	200
	600	232	179	184	204	168	90	230	204	139	232	235	236	236	233	217
	1000	236	183	196	205	216	92	236	201	147	236	238	236	236	212	186
AD	200	25.9	3.9	1.5	13.1	4.2	-4.7	16.5	-0.05	3.9	13.6	-9.6	1.3	34.8	-1.5	3.8
	600	26.4	5.7	3.2	14.9	5.8	-4.9	17.5	0.6	3.9	21	-2.6	2.1	37.2	-0.7	4.2
	1000	27.7	5.1	6.4	14.5	6.3	-4.6	18.6	0.06	4.3	25.6	0.77	1.6	41.4	3.7	4.2
SC	200	1.35	1.06	1.01	1.15	1.05	0.95	1.19	0.97	1.04	1.15	0.88	1.04	1.52	0.99	1.03
	600	1.4	1.1	1	1.2	1	0.95	1.2	0.97	1.04	1.3	0.96	1.1	1.6	1	1
	1000	1.39	1.07	1.06	1.17	1.08	0.95	1.22	0.97	1.05	1.34	0.98	1.05	1.68	1.04	1.03

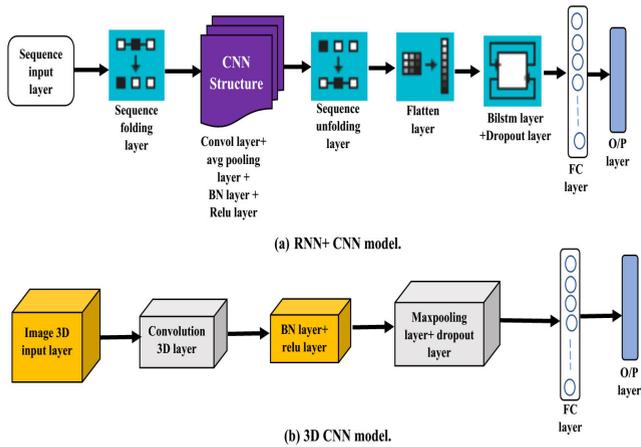


FIGURE 10. Models for 3D video classification and prediction.

of size 5×5 to attain the feature maps of the image. In the average pooling layer, we adopt a pooling size of 2×2 , and a stride of 2 to diminish the computational requirements. The nonlinear Rectified Linear Unit (ReLU) activation function is employed with the convolutional layers to permit contribution of other layers to the learning task, and the dropout layer is utilized to diminish the network overfitting probability. The batch normalization layer is used to accelerate the CNN training. The sequence unfolding layer and the flattening layer are used to restore the sequence structure and restructure the output to vector sequences, respectively. The BiLSTM

layer is used to learn bidirectional long-term dependencies between time steps of the sequence data and allow classification. Consequently, different numbers of nodes in the fully-connected layer according to the number of classes are correlated with the nodes in the average pooling layer [12], [13]. Finally, the detection of OAM states is performed through the SoftMax classifier.

B. 3D-CNN MODEL

This model consists of 9 layers: image 3D input layer, convolution 3D layer, max pooling 3D layer, ReLU layer, dropout layer, batch normalization layer, single fully-connected layer, and output layer. In 3D input layer, we input 3D images to the network, and after that, we apply the 3D convolution layer. This layer slips cubic shape convolution filters to the 3D input and then convolves the input with the filters by moving the filters along the input horizontally, vertically, and alongside the depth. The down-sampling process is performed using a 3D max pooling layer by splitting three-dimensional input into cuboidal pooling sections and computing the maximum of each region [11], [12].

VI. SIMULATION RESULTS FOR THE TWO USED DL DETECTION TECHNIQUES

A. CLASSIFICATION OF DIFFERENT 3D VIDEO FRAMES

After construction of the two DL models, we use them to classify and predict the different OAM-SK images, the different

codes and OAM-SK states. Now, to get the best model with an optimum response, the minimization of the training bias is required to make the training and validation loss values nearly the same or the difference between them small. After using different bias reduction techniques in the training process, the optimal values and the best datasets can be selected and used after that for the training process. A sample of the obtained results using datasets of 3 classes with LDPC code is shown in Table 5. According to the values in this table, it is confirmed that the training loss is very close to the validation loss for the two deep learning techniques at 10^{-4} regularization factor, dropout of 0.5, and 3 dataset classes. The cross-validation method shows that the utilization of far classes gives better results than those with near classes.

In Table 6, we introduce a comparison between the two used classification techniques: CRNN and 3D CNN, using different numbers of iterations and coded datasets. The comparison is performed using 3D chaotic interleaver, and the encryption process is performed on every used 3D video frame. From the table, we notice that LDPC code has the best performance, lowest validation loss, and least computation time compared to other codes. RS and Turbo codes have nearly the same accuracy, but Turbo code takes less time than RS code. BCH and convolution codes have the least validation and mini-batch accuracy compared to other codes, but the BCH code takes longer running time than that of the convolution code. Accordingly, convolution code is preferred to BCH code. When using the 6 classes in the training process, the accuracy of the two proposed models is diminished, but it increases when reducing the number of

classes. Moreover, the accuracy has the best value when using classes far from each other such as classes 2, and 64 or classes 4, and 64. This is due to the common OAM states, when the classes are close to each other. As long as the classes are far away from each other, the accuracy has better values. This is illustrated in Table 6 for different codes (L:LDPC, T:Turbo, R:RS, B:BCH, C:Convolution) and different classes (C). It is clear from the table that the accuracy is the best for the two-class case. When the number classes increases, the classification accuracy decreases.

B. CLASSIFICATION MODEL PERFORMANCE

After evaluating the accuracy of proposed models on different datasets, we measure other performance metrics of the deep network to demonstrate the model performance. The results are shown in Table 7. They are calculated according to the equations used in [14]. This table gives a comparison between the different performance metrics of the deep network using different codes and datasets from 3D chaotic interleaved coded OAM-SK. It is clear from the table that LDPC code is more efficient than other codes, and it has the best values of precision, Recall, SP, AUC, and FScore. BCH and convolution codes have nearly the same worst measurements of precision, Recall, specificity (SP), AUC, and FScore.

Furthermore, in Table 7, we study the DL performance metrics for the two used classification techniques. From the table, we find that the utilization of 2 classes makes both techniques for different codes have nearly the same deep learning performance. Increasing the number of classes

TABLE 5. Comparison between the training and validation losses using different bias reduction techniques.

Techniques		Cross-validation			Regularization			Dropout			Classes			Neurons		
		1	2	3	10^{-2}	10^{-4}	10^{-6}	0.2	0.5	0.7	2	3	6	1	2	3
T-Loss	3DCNN	0.37	0.33	0.35	0.39	0.32	0.44	0.2	0.20	0.25	0.003	0.31	0.75	0.18	0.28	0.71
	RCNN	0.21	0.25	0.28	0.56	0.29	0.29	0.15	0.29	0.19	0.0002	0.29	0.69	0.19	0.29	0.24
V-loss	3DCNN	0.25	0.26	0.26	0.27	0.28	0.26	0.23	0.23	0.30	0.001	0.34	0.74	0.25	0.29	0.36
	RCNN	0.28	0.28	0.31	0.33	0.30	0.32	0.30	0.30	0.33	0.0003	0.30	0.76	0.26	0.26	0.27

TABLE 6. Comparison between the accuracy of different classes of OAM-SK using different codes.

Datasets		CNN+ RNN Model						3D CNN Model					
		2 Classes		3 Classes		6 Classes		2 Classes		3 Classes		6 Classes	
Epoch/ Iterations		100/600	500/3000	100/900	500/4500	100/1900	500/9500	100/600	500/3000	100/900	500/4500	100/1900	500/9500
Time (minute) with accelerator (GPU)	L	18 s	1.48	39 s	4.2	1.03	6.32	50 s	4.1	1.14	6.15	2.49	14.26
	T	18 s	1.47	29 s	3.12	1.15	7.19	1.07	5.37	1.31	7.24	3.25	18.44
	R	20 s	1.59	28 s	2.54	1.05	6.35	51 s	4.14	1.17	6.30	2.5	14.33
	C	20 s	1.56	27 s	2.52	1.03	6.33	52 s	4.13	1.19	6.26	2.52	14.40
	B	21 s	1.57	32 s	3.17	1.03	6.33	1.05	5.02	1.38	7.27	3.24	19.02
Mini- Batch accuracy	L	100	100	100	87.5	59.75	75	100	100	96.88	93.75	65.63	50
	T	100	93.75	87.5	78.13	43.75	43.75	100	100	75	75	37.5	56.25
	R	100	100	84.38	87.5	59.38	59.38	100	100	93.75	87.5	53.13	50
	C	100	100	84.38	75	56.25	50	100	100	68.75	81.25	43.75	53.13
	B	100	100	78.13	84.38	56.25	62.5	100	100	81.25	87.5	40.63	62.5
Validation accuracy	L	100	100	90.37	90.37	70.37	70	100	100	91.85	91.85	67.41	68.89
	T	100	100	80.74	81.48	51.11	51.48	100	100	80	80.74	52.22	49.63
	R	100	100	87.41	87.41	61.85	61.85	100	100	85.19	85.19	62.59	63.33
	C	98.89	98.89	75.56	75.56	48.89	47.78	100	100	78.52	77.78	46.3	48.52
	B	100	100	86.67	83.7	57.78	57.41	100	100	82.22	82.96	54.44	56.67

TABLE 7. Comparison between a variety of performance metrics of deep learning for different classes of OAM and different codes.

Datasets		Accuracy		AUC		FScore		NPV		Precision		Recall		SP	
		1	2	1	2	1	2	1	2	1	2	1	2	1	2
2C	L	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	T	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	R	100	100	100	100	100	100	100	100	100	100	100	100	100	100
	C	98.8	100	98.9	100	98.8	100	98.8	100	98.9	100	98.9	100	98.9	100
	B	100	100	100	100	100	100	100	100	100	100	100	100	100	100
3C	L	93.5	94.5	93.8	94.1	90.2	91.7	95.1	95.9	90.3	91.8	92.1	93.4	95.6	96.3
	T	86	87.1	86.7	86.6	78.5	80.3	89.6	90.3	79.2	80.7	82.5	82.1	90.9	91
	R	91.6	90.1	92.7	91.8	86	84.4	93.7	92.5	87.4	85.1	90.8	89.7	94.7	93.9
	C	81.2	84.2	79.3	84.7	69	73.2	85.9	88.1	71.8	76.3	70.9	79	87.8	90.5
	B	89.1	88.1	88.2	87	83.3	81.3	91.8	91.1	83.7	82.2	84.11	82.3	92.3	91.8
6C	L	90	89.6	91.2	90.7	67.8	66.9	94	93.8	70.3	69.2	87.5	86.9	94.9	94.4
	T	83.6	83.4	72.9	71.3	49.7	47.8	90.3	90	51.4	50.3	55	51.9	90.7	90.6
	R	87.3	87.8	89.7	90.8	60	62.4	92	92.7	62.2	63.7	86.1	88	93.2	93.5
	C	82.2	82.4	72.3	74.9	43.8	47.4	89.3	89.4	46.6	47.4	54.7	59.8	90	90
	B	85.9	85.6	76	76.8	56.2	55.9	91.5	91.4	57.7	57	60.15	61.9	91	91.8

TABLE 8. Comparative study between the classification performances (%) of 2D-CNN and 3D-CNN using different codes.

Datasets		2 Classes		3 Classes		4 Classes	
Methods		3DCNN	2DCNN	3DCNN	2DCNN	3DCNN	2DCNN
Codes	LDPC	100	100	91.85	91.11	68.52	65.93
	Turbo	100	100	78.52	74.07	53.33	44.81
	RS	100	100	85.19	83.89	62.59	60.59
	BCH	100	98.89	80	76.3	57.04	49.63
	Conv	100	98.89	79.26	77.04	44.44	41.48

allows the performance with different codes to differ. With 3 and 6 classes, we find that LDPC code has the greatest values of accuracy, AUC, FScore, SP, Precision, and Recall compared to other codes.

One of the drawbacks of 2D CNN is handling the extraction of motion information that is performed outside the CNNs and consuming comparatively large time on GPUs. So, end-to-end methods for learning the representation of motion, such as 3D-CNN, can achieve faster and more accurate results. We present a 3D architecture for modeling the behavior and depicting motion in an easy way to be precise and to overcome real-time challenges. So, 3D-CNN technique is used to compute features instead of the 2D form in both spatial and temporal dimensions. The introduced model achieves a balance between the number of used frames and the number of operating parameters to achieve the best classification efficiency.

From Table 8, it is observed that the 2D-CNN and 3D-CNN techniques have nearly the same classification performance using only a two-class dataset. Increasing the number of classes in the dataset makes the 3D-CNN technique better than the 2D-CNN technique by nearly 0.74% to 10% according to the used code as obtained in [44].

C. PREDICTION MODEL PERFORMANCE

Here, different datasets are used to compute the prediction efficiency and attain the best dataset that is used to enhance the system performance.

In Fig. 11, it is claimed that the optimal expected responses for the different DL techniques are almost 100% for the different codes used employing only datasets of 2 class. We used the trial-and-error method to decide the best approach for dataset creation, as seen in the results obtained for the two deep learning techniques. It is shown that the farther the OAM states are moved from each other, the greater the effectiveness of prediction. When increasing the number of classes, it is shown that LDPC code has the best predictive performance due to its easy encoding and decoding processes. Next comes RS and Turbo codes, but Turbo code has a slightly worse predictive performance due to its high decoding complexity. Due to the complicated decoding performance of convolution and BCH codes, they have the worst predictive performance. Also, it is shown from the figure that OAM-SK states with $N = 2, 16,$ and 64 states have the best expected performance compared to the other states.

In Fig. 12, it is shown that when using datasets of different classes represented in codes, the performance of 3D CNN

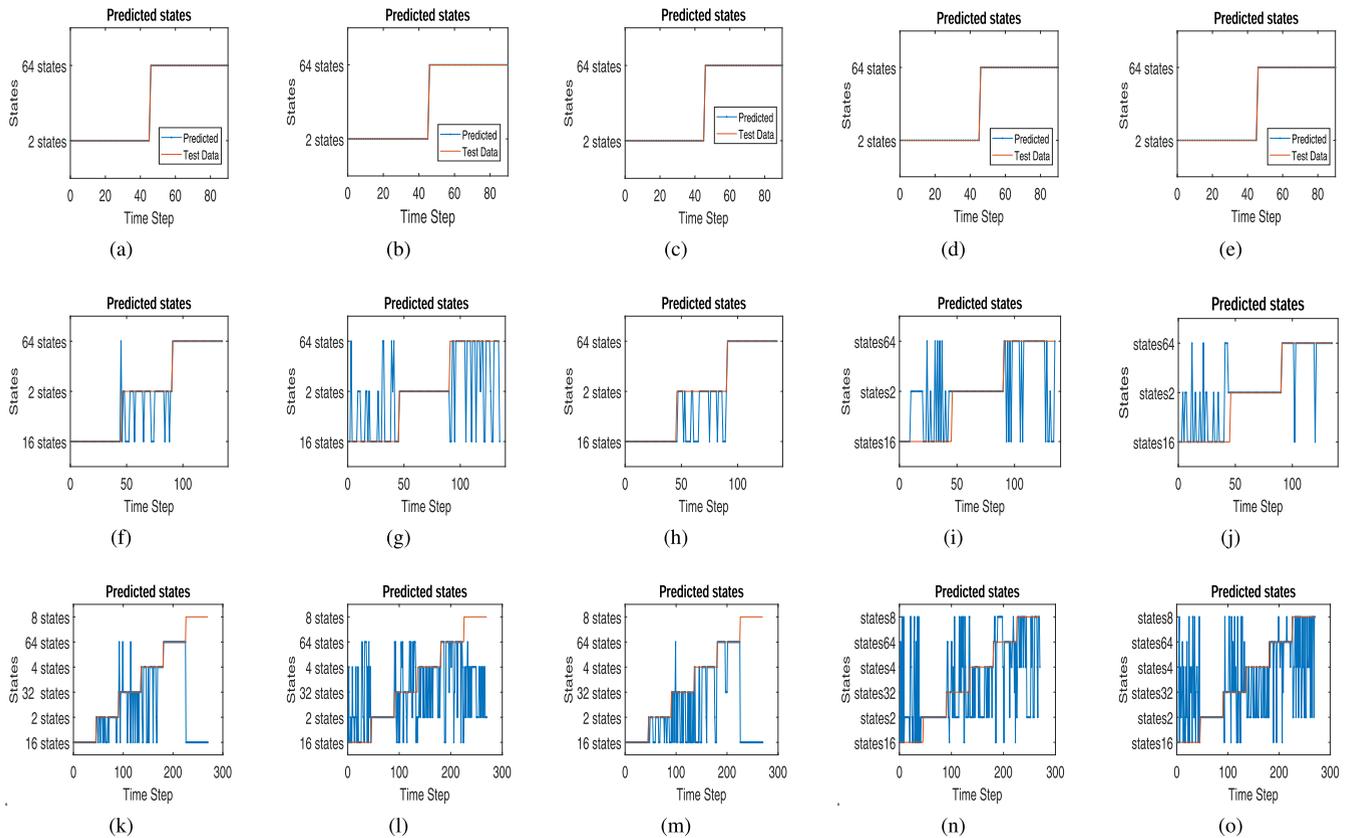


FIGURE 11. The optimum predicted responses of the two DL models for 3D chaotic interleaved video frames using datasets of different OAM state classes: (a), (f) and (k) LDPC code; (b), (g) and (l) Turbo code; (c), (h) and (m) RS code; (d), (i) and (n) Convolution code; (e), (j) and (o) BCH code.

model is more efficient than RNN + CNN model for the different classes. Also, LDPC code has the best predictive performance followed by RS and Turbo codes, and this occurs for the two used models.

Utilization of datasets of two classes makes the 3D CNN model more efficient compared to the RNN + CNN model. On the other hand, increasing the number of video frames makes the 3D CNN performance worse compared to that of the RNN + CNN as the final feature representation is obtained by combining information from all channels. In addition, increasing the time step values makes the RNN + CNN worse as shown in Fig. 13. Applying 4 video frames makes the prediction accuracy of the two used models the same. Due to the low speed of Newspaperfin video frames, it has the best predictive performance compared to the other used video frames. The Sharkfin comes after with moderate speed, and finally the fast-moving Dancerfin and Poznanstreetfin video frames with nearly the same performance.

From Table 9, it is shown that the prediction accuracy after using different datasets of different class number reaches 100%, when using datasets of 2 classes for the two used models. Increasing the number of classes reduces the prediction accuracy for the two models from 8% to 30%. After applying datasets of 3 classes, it is concluded that using

LDPC code makes the 3D CNN model more efficient than the CRNN model. For a dataset of 6 classes, the performance of the two models gets comparatively worse, but the CRNN model is a little superior to other models, when using LDPC code. For different VFs, it is clear that the prediction accuracy of the two used models is greatly affected by changing the used number of VFs. The utilization of two VFs makes the 3D CNN model more effective than the RNN + CNN model by about 5%. Increasing the number of VFs reduce the prediction accuracy of the two models, but makes the performance of them worse until reaching the same performance achieved when using 4 video frames. In the case of different codes, it is realized that the performance of 3D CNN model is superior to the RNN + CNN model for the different used datasets. Also, the prediction accuracy of the 3D CNN model reaches 100% using datasets of 2 and 3 classes.

To explain the usage of the RNN, a comparison has been performed between CNN and RNN using different datasets and different codes. From Table 10, it is shown that the presence of RNN with CNN enhances the system performance for different datasets. The classification performance for the different datasets using RNN + CNN is better than that using only CNN by nearly 2 : 12% according to the used code.

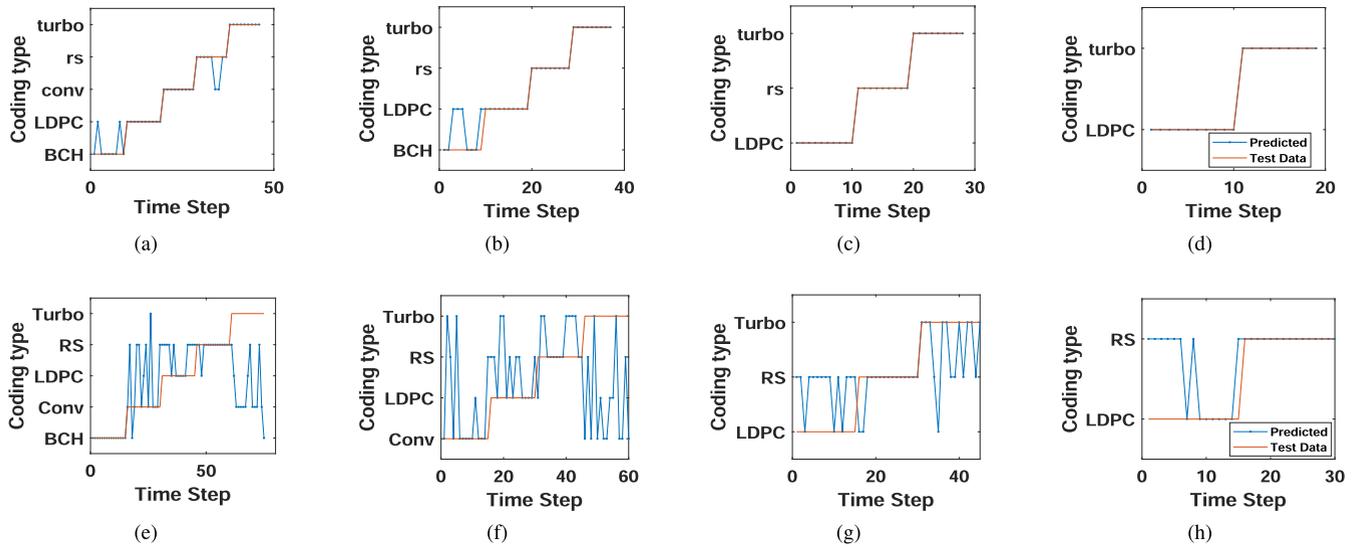


FIGURE 12. The predicted responses of 3D chaotic interleaved video frames using datasets of different codes: (a), (b), (c) and (d) 3D CNN model; (e), (f), (g), and (h) CRNN model.

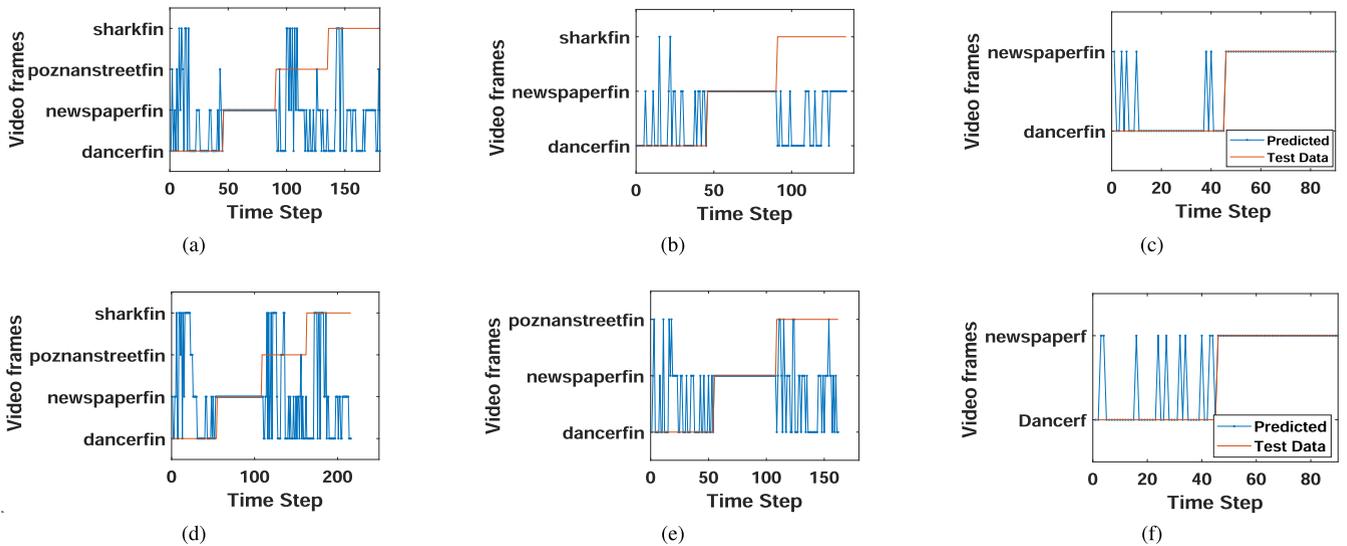


FIGURE 13. The predicted responses of 3D coded chaotic interleaved for N-OAM-SK model using datasets of different video frames: (a), (b), and (c) 3D CNN model; (d), (e), and (f) RNN + CNN model.

TABLE 9. The prediction accuracy (%) of the two applied DL models using different methodologies to construct different datasets via the turbulence N-OAM-SK model.

Datasets	Classes Number(CN)			Video Frames (VF)			Codes (C)			
	2 CN	3 CN	6 CN	2 VF	3 VF	4 VF	2 C	3 C	4 C	5 C
3D-CNN	100	92	69	93	56	44	100	100	90	91
CRNN	100	90	70	89	59	44	73	56	45	56

As shown from Table 11, a comparison is introduced between the two presented DL models with a variety of traditional image and video classification models. Our models are performed on datasets that come from chaotic interleaved N-OAM-SK in different scenarios. In [51], the 3D CSN model is used to factorize 3D convolutions by separating channel interactions and spatio-temporal interactions, result-

ing in greater accuracy of nearly 80% and lower computational costs compared to those of 3D convolutions. In [52], an efficient video classification technique with recurrent models was used to reduce the computation time for video classification using the idea of distillation. It is accomplished by firstly training the whole network, which gives a representation of all video frames and then training a

TABLE 10. The classification accuracy (%) comparison between CNN and CNN + RNN models.

Datasets		2 Classes		3 Classes		6 Classes	
Methods		2DCNN	CRNN	2DCNN	CRNN	2DCNN	CRNN
Codes	LDPC	100	100	91.11	93.5	65.93	70
	Turbo	100	100	74.07	86	44.81	51.85
	RS	100	100	83.89	91.6	60.59	61.52
	BCH	98.89	98.89	76.3	81.2	49.63	58.52
	Conv	98.89	100	77.04	89.1	41.48	49.63

TABLE 11. Comparison between the classification accuracy (%) of the introduced models and other traditional models.

Techniques	SVM	FDNN	CSN	CPPF	GCNN	SCNN	DFFN	OTAM	3DCNN	CRNN	CNN [50]	CNN [51]
Accuracy	77	85	83	81	83	82	85	86	96	95	86.7	94

network with some video frames only. It can reduce the computation time by nearly 30% compared to those of the two presented models that enhance the classification performance by almost 67.64% and 36.93% for CNN + RNN and 3D CNN models, respectively. From [53], it is clear that FDNN classification accuracy ranges from 77% to 98% by selected hyperparameters, and it is superior to DNN and RF classification accuracy. Due to the difficulty in collecting and clarifying large-scale video data in most DL models, a method called OTAM was used in this case as shown in [54]. It dynamically aligns two video sequences, while preserving the temporal ordering, and it is directly optimized for the few-shot learning objective in an end-to-end fashion. It achieves a classification accuracy of nearly 73 : 85.8% compared to 53 : 74% for the matching network that uses a Bi-LSTM. In [55], the classification accuracy obtained by various methods indicates that the DL-based methods outperform non-DL methods. The DFFN, which incorporates residual learning and feature fusion, achieves the best results in classification. The SVM and CNN-PPF depend only on spectral features during the classification. The rest of methods, including the 3D-CNN, Gabor-CNN, S-CNN, and DFFN belong to classification methods based on spectral-spatial features. In [56], the UrbanSound8K dataset was used to train and analyze the CNN classification performance. The training is performed using spectrogram images in an end-to-end fashion and the experimental results show that the proposed CNN model achieves a classification accuracy of 86.70%. Deeper networks can be trained with a novel architecture known as the convolutional highway unit [57]. The unit architecture is formed by modified convolutional highway layers, a max pool layer, and a dropout layer. According to the experimental results, it can achieve a 94% classification accuracy, when the training data is reduced to 30% of its original size.

The introduced DL models outperform all previous ones as indicated in Table 11.

The reasons can be summarized as follows:

- 1) In DFFN, the best feature fusion mechanism requires a hand-crafted framework with a lot of experiments. Therefore, the performance is worse, when the model complexity and size of the dataset increase [55].

- 2) Both CNN-PPF and S-CNN concentrate on analyzing the correlation between samples rather than pixel-wise semantic information [55].
- 3) CNN-PPF uses a CNN to retrieve the pixel-pair features. However, because the convolution operation is primarily performed in the spectral domain, spatial information is not taken into account for the CNN-PPF [55].
- 4) S-CNN, on the other hand, uses a two-branch CNN to extract spectral-spatial features at the same time. However, the computational cost of such method may be prohibitively high due to the high-dimension vector in Euclidean space [55].
- 5) In G-CNN, Gabor filtering is used as a preprocessing technique to extract spatial features of hyperspectral images (HSIs).
- 6) The great spatial variation of spectral signatures, as well as the limited available training samples versus the high dimensionality of hyperspectral data, present two major challenges in HSI classification tasks [55].
- 7) The SVM was implemented in the LIB-SVM library as a spectral feature-based tool, with a Gaussian kernel and fivefold cross validation. The performance is bad in the presence of noise or turbulence [55].
- 8) In CSN [51], when there are sufficient channel interactions in the network, the CSN-based factorization not only helps to reduce the computational cost, but it also increases the accuracy.
- 9) When signals are sparse and correlated, the FDNN classifier outperforms ordinary random forests or DNNs alone in classification. Furthermore, for random forests and DNNs, an improved performance could not be achieved by simply increasing model complexity [53].
- 10) Furthermore, OTAM is completely differentiable, allowing an end-to-end training model to optimize the few-shot learning target. As a result, the model has better ability to use long-term temporal information. It automatically aligns two video sequences, while maintaining temporal ordering [54].

The ability to generalize is an essential metric for assessing the quality of a CNN model, and it is a hot and difficult research subject. The perfect classification results show that

the well-trained CNN generalization ability is very good. When we train a CNN model using images created by known computer graphics rendering techniques, and then evaluate the model on images generated by unknown rendering techniques, the generalization ability is tested [58]. In this work, the different datasets come from applying different atmospheric turbulence conditions, and then the optimum datasets are chosen. These optimum datasets are used to study the classification and prediction performance of the two proposed models. The classification performance of the two proposed models is evaluated and compared with those of other traditional ones as shown in Table 11. From this table, it is indicated that the two proposed models outperform the other traditional ones. The efficient classification results show that the generalization of the two proposed models is possible.

VII. CONCLUSION AND FUTURE WORK

We have investigated encrypted coded VF transmission through 16-OAM-SK-FSO model with different AT parameters. It has been proved that the LDPC coded Newspaper VFs have the lowest BER compared to those of other coded VFs with the aid of 2D and 3D chaotic interleavers. Employing 3D CI for VF transmission through N -OAM-SK-FSO model improves the PSNR by nearly 3 dB compared to that of the 2D CI scenario. Applying a 2-class datasets for video classification and prediction makes the two DL models have nearly the same accuracy, but the situation is altered, when using datasets of larger classes. The implementation of the two models indicates that they have the same Accuracy, AUC, FScore, and other performance metrics of deep networks for different datasets of different OAM states. Diminishing the number of classes makes LDPC and Turbo codes with the 3D CNN model better than those with the CRNN model. The utilization of GPU in classification enhances the classification performance by nearly 52.285% for CRNN and 3D CNN models. It has been proved from the obtained results that 3D CNN and CRNN models have superior performance compared to those of the 2D CNN and the other traditional deep learning models by nearly 10 : 18%.

As a future work, to further improve the model efficiency and resolve the defects of the OAM-SK optical system, it would be important to consider OAM performance with other different and effectual DL techniques. Although our research was limited to the use of SISO OAM, all the developed techniques are applicable to the more efficient MIMO OAM. We can merge higher order modulation techniques with OAM in addition to further theoretical analysis. Moreover, OAM-based orthogonal and non-orthogonal multiple access (OAM/NOMA) system can be studied. Furthermore, we are interested in studying various DL techniques with other distinct security algorithms. With OAM-DL systems, we can also study different encryption/decryption schemes to enhance the efficiency and reliability of the systems. Recently, there have been many works that study the performance of OAM mode recognition using LG beams. As a future work, we are going to use Bessel Gaussian beams, and

hyper-geometric Gaussian beams to study the performance of OAM mode recognition. The performance of OAM under different parameters such as amplified spontaneous emission noise, pointing error, multiplicative slow fading, and geometric spreading will be studied and analyzed to further improve the efficiency of the model.

ACKNOWLEDGMENT

The authors would like to acknowledge the support received from Taif University Researchers Supporting Project Number (TURSP-2020/147), Taif university, Taif, Saudi Arabia.

REFERENCES

- [1] T. A. Tsiftsis, H. G. Sandalidis, G. K. Karagiannidis, and M. Uysal, "Optical wireless links with spatial diversity over strong atmospheric turbulence channels," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 951–957, Feb. 2009.
- [2] W. Cheng, W. Zhang, H. Jing, S. Gao, and H. Zhang, "Orbital angular momentum for wireless communications," *IEEE Wireless Commun.*, vol. 26, no. 1, pp. 100–107, Feb. 2019.
- [3] B. Yan, Z. Lu, J. Hu, T. Xu, H. Zhang, W. Lin, Y. Yue, H. Liu, and B. Liu, "Orbital angular momentum (OAM) carried by asymmetric vortex beams for wireless communications: Theory, experiment and current challenges," *IEEE J. Sel. Topics Quantum Electron.*, vol. 27, no. 2, pp. 1–10, Mar. 2021.
- [4] C. Kai, P. Huang, F. Shen, H. Zhou, and Z. Guo, "Orbital angular momentum shift keying based optical communication system," *IEEE Photon. J.*, vol. 9, no. 2, pp. 1–10, Apr. 2017.
- [5] W. Wang, P. Wang, T. Cao, H. Tian, Y. Zhang, and L. Guo, "Performance investigation of underwater wireless optical communication system using M-ary OAMSK modulation over oceanic turbulence," *IEEE Photon. J.*, vol. 9, no. 5, pp. 1–15, Oct. 2017.
- [6] Y. Li, C. Wang, X. Zhang, J. Liu, Y. He, J. Xiao, Y. Gao, S. Chen, X. Zhang, and D. Fan, "Optical orbital angular momentum shift-keying communication using direct demodulation," *IEEE Access*, vol. 7, pp. 103433–103442, 2019.
- [7] S.-C. Chu, Y.-T. Chen, K.-F. Tsai, and K. Otsuka, "Generation of high-order Hermite-Gaussian modes in end-pumped solid-state lasers for square vortex array laser beam generation," *Opt. Exp.*, vol. 20, no. 7, pp. 7128–7141, 2012.
- [8] M. Padgett, J. Courtial, and L. Allen, "Light's orbital angular momentum," *Phys. Today*, vol. 57, no. 5, pp. 35–40, 2004.
- [9] F. Dufaux, B. Pesquet-Popescu, and M. Cagnazzo, *Emerging Technologies for 3D Video: Creation, Coding, Transmission and Rendering*. Hoboken, NJ, USA: Wiley, 2013.
- [10] T. Zhang and S. Mao, "Joint power and channel resource optimization in soft multi-view video delivery," *IEEE Access*, vol. 7, pp. 148084–148097, 2019.
- [11] K. Guo, Y. Hu, Z. Qian, H. Liu, K. Zhang, Y. Sun, J. Gao, and B. Yin, "Optimized graph convolution recurrent neural network for traffic prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 22, no. 2, pp. 1138–1149, Feb. 2021.
- [12] Z. Wu, T. Yao, Y. Fu, and Y.-G. Jiang, "Deep learning for video classification and captioning," in *Frontiers of Multimedia Research*. New York, NY, USA: ACM, Dec. 2017, pp. 3–29. [Online]. Available: https://dl.acm.org/doi/abs/10.1145/3122865.3122867?casa_token=1LA2pVPZeQgAAAAA%3A1fj7zP7kN_6KL1weLX00t1_isGUxI3ICJk7dngO2bdYbpyumw9jv5dQ0tVfazeQ0I69OfUzU0x,doi:10.1145/3122865.3122867.
- [13] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2013.
- [14] Z. Wang, M. I. Dedo, K. Guo, K. Zhou, F. Shen, Y. Sun, S. Liu, and Z. Guo, "Efficient recognition of the propagated orbital angular momentum modes in turbulences with the convolutional neural network," *IEEE Photon. J.*, vol. 11, no. 3, pp. 1–14, Jun. 2019.
- [15] M. I. Dedo, Z. Wang, K. Guo, and Z. Guo, "OAM mode recognition based on joint scheme of combining the Gerchberg-Saxton (GS) algorithm and convolutional neural network (CNN)," *Opt. Commun.*, vol. 456, Feb. 2020, Art. no. 124696.

- [16] J. Li, M. Zhang, and D. Wang, "Adaptive demodulator using machine learning for orbital angular momentum shift keying," *IEEE Photon. Technol. Lett.*, vol. 29, no. 17, pp. 1455–1458, Sep. 1, 2017.
- [17] Z. Wang and Z. Guo, "Adaptive demodulation technique for efficiently detecting orbital angular momentum (OAM) modes based on the improved convolutional neural network," *IEEE Access*, vol. 7, pp. 163633–163643, 2019.
- [18] E. Knutson, S. Lohani, O. Danaci, S. D. Huver, and R. T. Glasser, "Deep learning as a tool to distinguish between high orbital angular momentum optical modes," *Proc. SPIE*, vol. 9970, Sep. 2016, Art. no. 997013.
- [19] Q. Tian, Z. Li, K. Hu, L. Zhu, X. Pan, Q. Zhang, Y. Wang, F. Tian, X. Yin, and X. Xin, "Turbo-coded 16-ary OAM shift keying FSO communication system combining the CNN-based adaptive demodulator," *Opt. Exp.*, vol. 26, no. 21, pp. 27849–27864, 2018.
- [20] H. Song, H. Song, R. Zhang, K. Manukyan, L. Li, Z. Zhao, K. Pang, C. Liu, A. Almainan, R. Bock, B. Lynn, M. Tur, and A. E. Willner, "Experimental mitigation of atmospheric turbulence effect using pre-signal combining for uni- and bi-directional free-space optical links with two 100-gbit/s OAM-multiplexed channels," *J. Lightw. Technol.*, vol. 38, no. 1, pp. 82–89, Jan. 2020.
- [21] J. Liu, P. Wang, X. Zhang, Y. He, X. Zhou, H. Ye, Y. Li, S. Xu, S. Chen, and D. Fan, "Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication," *Opt. Exp.*, vol. 27, no. 12, pp. 16671–16688, 2019.
- [22] C. Chen, H. Yang, Y. Lou, and S. Tong, "Changes in orbital-angular-momentum modes of a propagated vortex Gaussian beam through weak-to-strong atmospheric turbulence," *Opt. Exp.*, vol. 24, no. 7, pp. 6959–6975, Apr. 2016.
- [23] S. A. El-Meadawy, H. M. Shalaby, N. A. Ismail, F. E. A. El-Samie, and A. E. Farghal, "Free-space 16-ary orbital angular momentum coded optical communication system based on chaotic interleaving and convolutional neural networks," *Appl. Opt.*, vol. 59, no. 23, pp. 6966–6976, 2020.
- [24] D. Mamadou, F. Shen, M. Dedo, Q. Zhou, K. Guo, and Z. Guo, "High-efficiency sorting and measurement of orbital angular momentum modes based on the March–Zehnder interferometer and complex phase gratings," *Meas. Sci. Technol.*, vol. 30, no. 7, Jul. 2019, Art. no. 075201.
- [25] Z. Feng, X. Wang, M. I. Dedo, K. Guo, F. Shen, C. Kai, and Z. Guo, "High-density orbital angular momentum mode analyzer based on the mode converters combining with the modified Mach–Zehnder interferometer," *Opt. Commun.*, vol. 435, pp. 441–448, Mar. 2019.
- [26] C. Kai, Z. Feng, M. I. Dedo, P. Huang, K. Guo, F. Shen, J. Gao, and Z. Guo, "The performances of different OAM encoding systems," *Opt. Commun.*, vol. 430, pp. 151–157, Jan. 2019.
- [27] Z. Guo, Z. Wang, M. I. Dedo, and K. Guo, "The orbital angular momentum encoding system with radial indices of Laguerre–Gaussian beam," *IEEE Photon. J.*, vol. 10, no. 5, pp. 1–11, Oct. 2018.
- [28] H. S. Khallaf, A. E. Morra, A. E. Elfiqi, H. M. Shalaby, and S. Hranilovic, "Hybrid two-level MPPM–MDPSK modulation for high-speed optical communication networks," *Appl. Opt.*, vol. 58, no. 36, pp. 9757–9767, 2019.
- [29] X. Fang, H. Ren, and M. Gu, "Orbital angular momentum holography for high-security encryption," *Nature Photon.*, vol. 14, no. 2, pp. 102–108, Feb. 2020.
- [30] Z. Dong, Y. Zhang, H. Li, R. Tao, C. Gu, P. Yao, Q. Zhan, and L. Xu, "Generation of stable orbital angular momentum beams with an all-polarization-maintaining fiber structure," *Opt. Exp.*, vol. 28, no. 7, pp. 9988–9995, 2020.
- [31] M. Sharma, F. Amir Khan, S. K. Mishra, D. Sengupta, Y. Messaddeq, F. Blanchard, and B. Ung, "Transmission of orbital angular momentum and cylindrical vector beams in a large-bandwidth annular core photonic crystal fiber," *Fibers*, vol. 8, no. 4, p. 22, Apr. 2020.
- [32] W. Wang, P. Wang, L. Guo, W. Pang, W. Chen, A. Li, and M. Han, "Performance investigation of OAMSK modulated wireless optical system over turbulent ocean using convolutional neural networks," *J. Lightw. Technol.*, vol. 38, no. 7, pp. 1753–1765, Apr. 1, 2020.
- [33] X. Ding, G. Feng, and S. Zhou, "Detection of phase distribution of vortex beams based on low frequency heterodyne interferometry with a common commercial CCD camera," *Appl. Phys. Lett.*, vol. 116, no. 3, Jan. 2020, Art. no. 031106.
- [34] Y. Song and Y.-S. Ho, "Unified depth intra coding for 3D video extension of HEVC," *Signal, Image Video Process.*, vol. 8, no. 6, pp. 1031–1037, Sep. 2014.
- [35] W. El-Shafai, S. El-Rabaie, M. M. El-Halawany, and F. E. A. El-Samie, "Encoder-independent decoder-dependent depth-assisted error concealment algorithm for wireless 3D video communication," *Multimedia Tools Appl.*, vol. 77, no. 11, pp. 13145–13172, Jun. 2018.
- [36] M. B. Hossain, M. T. Rahman, A. B. M. S. Rahman, and S. Islam, "A new approach of image encryption using 3D chaotic map to enhance security of multimedia component," in *Proc. Int. Conf. Informat., Electron. Vis. (ICIEV)*, May 2014, pp. 1–6.
- [37] B. Wang, P. Chen, Y. Fang, and F. C. M. Lau, "The design of vertical RS-CRC and LDPC code for ship-based satellite communications on-the-move," *IEEE Access*, vol. 7, pp. 44977–44986, 2019.
- [38] E. M. Amhoud, A. Trichili, B. S. Ooi, and M.-S. Alouini, "OAM mode selection and space–time coding for atmospheric turbulence mitigation in FSO communication," *IEEE Access*, vol. 7, pp. 88049–88057, 2019.
- [39] L. Allen, M. Padgett, and M. Babiker, "Vortex beams: The orbital angular momentum of light," in *Progress in Optics*, vol. 39. Amsterdam, The Netherlands: Elsevier, 1999, pp. 291–372.
- [40] S. M. Mohammadi, L. K. S. Daldorff, J. E. S. Bergman, R. L. Karlsson, B. Thide, K. Forozesh, T. D. Carozzi, and B. Isham, "Orbital angular momentum in radio—A system study," *IEEE Trans. Antennas Propag.*, vol. 58, no. 2, pp. 565–572, Feb. 2010.
- [41] E.-M. Amhoud, B. S. Ooi, and M.-S. Alouini, "A unified statistical model for atmospheric turbulence-induced fading in orbital angular momentum multiplexed FSO systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 888–900, Feb. 2020.
- [42] Y. Ren, Z. Wang, G. Xie, L. Li, A. J. Willner, Y. Cao, Z. Zhao, Y. Yan, N. Ahmed, N. Ashrafi, S. Ashrafi, R. Bock, M. Tur, and A. E. Willner, "Atmospheric turbulence mitigation in an OAM-based MIMO free-space optical link using spatial diversity combined with MIMO equalization," *Opt. Lett.*, vol. 41, no. 11, pp. 2406–2409, 2016.
- [43] Z. Ghassemloooy, W. Popoola, and S. Rajbhandari, *Optical Wireless Communications: System and Channel Modelling With MATLAB*. Boca Raton, FL, USA: CRC Press, 2019.
- [44] V. Argyriou, J. M. Del Rincón, B. Villarini, and A. Roche, *Image, Video 3D Data Registration*. Chichester, U.K.: Wiley, 2015.
- [45] Y. Zhang, P. Wang, T. Liu, L. Guo, Y. Li, and W. Wang, "Performance analysis of a ldpc coded oam-based uca fso system exploring linear equalization with channel estimation over atmospheric turbulence," *Opt. Exp.*, vol. 26, no. 17, pp. 22182–22196, 2018.
- [46] I. B. Djordjevic, "Multidimensional OAM-based secure high-speed wireless communications," *IEEE Access*, vol. 5, pp. 16416–16428, 2017.
- [47] I. B. Djordjevic, "Heterogeneous transparent optical networking based on coded OAM modulation," *IEEE Photon. J.*, vol. 3, no. 3, pp. 531–537, Jun. 2011.
- [48] I. B. Djordjevic, J. A. Anguita, and B. Vasic, "Error-correction coded orbital-angular-momentum modulation for FSO channels affected by turbulence," *J. Lightw. Technol.*, vol. 30, no. 17, pp. 2846–2852, Sep. 2012.
- [49] I. B. Djordjevic, "OAM-based hybrid free-space optical-terahertz multidimensional coded modulation and physical-layer security," *IEEE Photon. J.*, vol. 9, no. 4, pp. 1–12, Aug. 2017.
- [50] S. Huang and M. Safari, "Free-space optical communication impaired by angular fluctuations," *IEEE Trans. Wireless Commun.*, vol. 16, no. 11, pp. 7475–7487, Nov. 2017.
- [51] D. Tran, H. Wang, M. Feiszli, and L. Torresani, "Video classification with channel-separated convolutional networks," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 5552–5561.
- [52] S. Bhardwaj, M. Srinivasan, and M. M. Khapra, "Efficient video classification using fewer frames," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 354–363.
- [53] Y. Kong and T. Yu, "A deep neural network model using random forest to extract feature representation for gene expression data classification," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, Dec. 2018.
- [54] K. Cao, J. Ji, Z. Cao, C.-Y. Chang, and J. C. Nibbles, "Few-shot video classification via temporal alignment," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 10618–10627.
- [55] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019.
- [56] F. Demir, D. A. Abdullah, and A. Sengur, "A new deep CNN model for environmental sound classification," *IEEE Access*, vol. 8, pp. 66529–66537, 2020.
- [57] Z. Lin, K. Ji, M. Kang, X. Leng, and H. Zou, "Deep convolutional highway unit network for SAR target classification with limited labeled training data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 7, pp. 1091–1095, Jul. 2017.
- [58] T. Zhang and H. Tang, "Evaluating the generalization ability of convolutional neural networks for built-up area extraction in different cities of China," *Optoelectron. Lett.*, vol. 16, no. 1, pp. 52–58, Jan. 2020.

SHIMAA A. EL-MEADAWY received the B.Sc. and M.Sc. degrees in electrical engineering from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 2013 and 2017, respectively. She is currently pursuing the Ph.D. degree in electrical engineering. She is currently an Assistant Lecturer with the Department of electrical Engineering, Faculty of Electronic Engineering, Menoufia University. Her current research interests include optical communications, deep learning (DL) applications and computer vision, deep neural networks, information theory, elastic optical networks, image and video signal processing, efficient 2-D video/3-D multi-view video coding, multi-view video plus depth coding, 3-D multi-view video coding and transmission, quality of service and experience, digital communication techniques, speech processing, and security algorithms.



HOSSAM M. H. SHALABY (Senior Member, IEEE) was born in Giza, Egypt, in 1961. He received the B.S. and M.S. degrees from Alexandria University, Alexandria, Egypt, in 1983 and 1986, respectively, and the Ph.D. degree from the University of Maryland at College Park, in 1991, all in electrical engineering. In 1991, he joined the Department of Electrical Engineering, Alexandria University, where he was promoted to a Professor, in 2001. From 1996 to 1998,

he was with the Department of Electrical and Computer Engineering, International Islamic University Malaysia, and the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 1998 to 2001. From December 2000 to 2004, he was an Adjunct Professor with the Department of Electrical and Information Engineering, Faculty of Sciences and Engineering, Laval University, Quebec City, QC, Canada. From 2007 to 2010, he worked as a Consultant with SysDSoft Company, Alexandria. From 2010 to 2016, he was the Chair of the Department of ECE, E-JUST. Since 2017, he has been on leave from Alexandria University. He is currently a Professor with the Department of Electronics and Communications Engineering (ECE), School of Electronics, Communications, and Computer Engineering, Egypt-Japan University of Science and Technology (E-JUST), New Borg EL-Arab City, Egypt. His research interests include optical communications, silicon photonics, optical CDMA, and quantum information theory. Since 2020, he has been serving as a Student Branch Counselor for the IEEE E-JUST Student Branch.



NABIL A. ISMAIL received the Ph.D. degree in computer engineering from Durham University, in 1983. From August 2006 to August 2008, he was the Dean of the Faculty of Computers and Information, Menoufia University. He is currently a Professor of computer science and engineering, Faculty of Electronic Engineering, Menoufia University. His main research interests include deep learning applications and computer vision, tomography, computer security, computer architecture,

elliptic curve cryptography, processor design, light-power smart devices, security applications, EIT algorithms, and multi-core/many-core parallel programming.



AHMED E. A. FARGHAL received the B.Sc. (Hons.) and M.Sc. degrees from Menoufia University, Menoufia, Egypt, and the Ph.D. degree from the Egypt-Japan University of Science and Technology, Egypt, all in electrical engineering, in 2006, 2011, and 2015, respectively. In 2007, he joined the Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, where he was promoted to the position of a

Lecturer Assistant, in 2011. From November 2014 to July 2015, he joined Kyushu University, Fukuoka, Japan, as a Special Research Student. He was with the Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, where he

was an Assistant Professor. He is currently with the Department of Electrical Engineering, Faculty of Engineering, Sohag University, Sohag, Egypt. His research interests include optical CDMA, all-optical networks, QoS provisioning in optical networks, elastic optical networking, green optical networks, and nano-optoelectronic devices.



FATHI E. ABD EL-SAMIE received the B.Sc. (Hons.), M.Sc., and Ph.D. degrees from Menoufia University, Menouf, Egypt, in 1998, 2001, and 2005, respectively. Since 2005, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. His current research interests include image enhancement, image restoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications. He was a recipient of the Most Cited Paper Award from the *Digital Signal Processing* journal, in 2008.



MOHAMMED ABD-ELNABY received the B.S., M.S., and Ph.D. degrees in electronic engineering from Menoufia University, Menouf, Egypt, in 2000, 2004, and 2010, respectively. Since 2010, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. Since 2015, he has also been working as an Associate Professor with the Department of Electronics and Electrical Communication, Faculty of Electronic Engineering, Menoufia University. He has coauthored about 69 articles in international journals and conference proceedings. His research interests include wireless resource management, MAC protocols, cognitive radio, cooperative communication, the IoT, 5G communication, NOMA, and D2D communication.



WALID EL-SHAFAI was born in Alexandria, Egypt. He received the B.Sc. degree (Hons.) in electronics and electrical communication engineering from the Faculty of Electronic Engineering (FEE), Menoufia University, Menouf, Egypt, in 2008, the M.Sc. degree from the Egypt-Japan University of Science and Technology (E-JUST), in 2012, and the Ph.D. degree from the Faculty of Electronic Engineering, Menoufia University, in 2019. Since January 2021, he has been joined

as a Postdoctoral Research Fellow with the Security Engineering Laboratory (SEL), Prince Sultan University (PSU), Riyadh, Saudi Arabia. He is currently working as a Lecturer and an Assistant Professor with the Department of Electronics and Communication Engineering (ECE), FEE, Menoufia University. His research interests include wireless mobile and multimedia communications systems, image and video signal processing, efficient 2-D video/3-D multi-view video coding, multi-view video plus depth coding, 3-D multi-view video coding and transmission, quality of service and experience, digital communication techniques, cognitive radio networks, adaptive filters design, 3-D video watermarking, steganography, and encryption, error resilience and concealment algorithms for H.264/AVC, H.264/MVC, and H.265/HEVC video codecs standards, cognitive cryptography, medical image processing, speech processing, security algorithms, software defined networks, the Internet of Things, medical diagnoses applications, FPGA implementations for signal processing algorithms and communication systems, cancellable biometrics and pattern recognition, image and video magnification, artificial intelligence for signal processing algorithms and communication systems, modulation identification and classification, image and video super-resolution and denoising, cybersecurity applications, deep learning in signal processing, and communication systems applications. He serves as a reviewer for several international journals.

...