The Institution of
Engineering and Technology | WILEY

ORIGINAL RESEARCH PAPER

Low complexity deep learning algorithms for compensating atmospheric turbulence in the free space optical communication system

Mohammad Ali Amirabadi [ID] | Mohammad Hossein Kahaei | S. Alireza Nezamalhosseni

School of Electrical Engineering, Iran University of Science and Technology, Tehran, Iran

Correspondence

S. Alireza. Nezamalhosseni, School of Electrical Engineering, Iran University of Science and Technology (IUST), Tehran 1684613114, Iran.
Email: nezam@iust.ac.ir

Abstract

One of the main barriers of free space optical (FSO) communication systems is atmospheric turbulence. Various processing techniques at the transmitter, receiver, and transceiver sides are available for addressing this issue; however, they have either high complexity or low performance. Considering this problem, in this study, deep learning (DL) is deployed at the transmitter, receiver, and transceiver sides of an FSO system for constellation shaping, detection, and joint constellation-shaping detection, respectively. Furthermore, the proposed DL-based structures are deployed in an FSO-multi-input multi-output (MIMO) system. As the first investigation over DL for the FSO-MIMO system, different combining schemes including the maximum ratio combiner, equal gain combiner, and the selection combiner are considered. Considering a wide range of atmospheric turbulence, from the weak to the strong regime, the performance of the proposed structures are compared with that of the maximum likelihood (ML) detection. To the best of the authors' knowledge, the main contributions and novelties of this work include considering transmitter learning in the FSO system, designing low complexity DL structures for FSO system applications, and providing complexity analysis for the proposed DL algorithms. The results indicate that the proposed DL-based FSO systems achieve the optimum performance with lower complexity compared with the state-of-the-art conventional FSO systems. For instance, the proposed DL-based detector is almost 2, 3, and 7.5 times faster than the ML detector for modulation orders of 16, 64, and 256, respectively.

KEYWORDS

atmospheric turbulence, free-space optical communication, signal detection

1 | INTRODUCTION

Recently, the free space optical (FSO) communication system has been taken into consideration in research studies. This is because of its high data rate, bandwidth and security in an unlicensed spectrum as well as its easy and low-cost installation [1–3]. In contrast with these advantages, high sensitivity of the FSO system to weather conditions and atmospheric turbulences limits its practical applications. The free space optical communication system is the main competitor against the conventional radio frequency (RF) communication system and is widely applied in space and ground communication links. Compared with the conventional RF system, the FSO system has a huge un-licenced bandwidth with a high data rate, high security, low cost, and low power consumption as well as favourable performance and easy installation. Furthermore, the FSO system can be used at the bottleneck and last-mile applications and is a good candidate for the back-haul network of the next generation communication network [1].

Although the FSO system enjoys many advantages over the conventional RF system, it has limited practical applications due to its low performance in long-range links. The main performance degradation factor in long-range links is

atmospheric turbulence [4]. The inhomogeneous disturbance of temperature and pressure over the atmosphere results in random changes of the refractive index. This induces phase disturbances and leads to intensity fluctuations, beam wandering and beam broadening of the FSO signal. The random fluctuations in the FSO signal intensity is called scintillation and is the most important effect of atmospheric turbulence. Several solutions have been proposed to mitigate the atmospheric turbulence effect. One of the promising approaches is the multi-input multi-output (MIMO) technique [5, 6]. MIMO is a well-established technique known for better performance, higher capacity, and increased coverage area compared to the single-input single-output (SISO) technique. Although maximum likelihood (ML) provides an optimal detection scheme, its complexity is considerable. Even though there are several sub-optimal detectors with lower complexity, they have reduced performance compared with the ML detector [7].

1.1 | Related works

Recently, machine learning techniques have been successfully applied as a powerful detector for different optical communication (OC) systems [8] including visible light communication [9, 10], fibre OC [11–14], and FSO communication [15]. In the context of FSO communication, many algorithms are used to detect the received signal after propagating through atmospheric turbulence. Depending on the deployed algorithm, these works can be divided into three categories; classical machine learning-based methods [16–19], convolutional neural network (CNN)-based methods [20–32], and deep neural network (DNN)-based methods [33–40]. Table 1 shows the main differences between these works.

Considering the classical machine learning category, the support vector machine is proposed in Ref [16] as a detector in the FSO-SISO system with on–off keying (OOK) to combat the noise and scintillation effects. The decision tree for optical power estimation in a hybrid FSO/RF system with OOK and hard switching considering weather conditions is studied in Ref [17]. A backpropagation artificial neural network is applied in Ref [18] for distortion correction in a sensor-less adaptive optic system. The artificial neural network is also used in Ref [19] for mitigating the atmospheric turbulence effects at the beam level without the need for adaptive optical kits.

Considering the CNN category, the CNN is deployed in Ref [20, 21] for identification of the structured light intensity profiles without requiring mode sorters or mode demultiplexers and in Ref [22] for detection in the FSO-SISO system with OOK. Orbital angular momentum (OAM) encoding is a promising approach for enhancing the FSO system capacity [23]. However, atmospheric turbulence causes distortion of phase fronts of OAM beams and hinders the decoding of OAM modes. Different CNN structures are deployed in Ref [23, 24] for efficient decoding of the OAM modes. In Ref [25–27], the CNN is used as a demodulator for a turbo-coded OAM-shift keying (SK) FSO system at strong atmospheric turbulence. In Ref [28], the CNN is used for joint atmospheric turbulence detection and demodulation for an OAM FSO system. Compared to previous approaches using the self-organising mapping, this method achieved a higher accuracy in atmospheric turbulence detection. In Ref [29], three-dimensional chaotic interleaved multi-coded video frames are transmitted using the OAM-SK FSO system to have reliable video communication. To tackle the defects of the OAM-SK FSO system, a three-dimensional CNN is proposed to decode OAM modes. In Ref [30], the authors used CNN as a demodulator for the OAM-SK FSO system considering the effects of pointing error and limited receiving aperture. The authors of Ref [31] presented a double stage system to receive high fidelity image data wherein first, a multi-CNN demodulator detects the atmospheric turbulence strength, and then a CNN demodulates the incident OAM modes. To reduce the atmospheric turbulence-induced power loss, an AlexNet-based CNN is deployed in Ref [32] for wavefront aberration compensation.

Considering the DNN category, an autoencoder for the FSO-SISO system with OOK is presented in Ref [33] considering a log-normal turbulence channel and the perfect channel state information (CSI). In Ref [34], an autoencoder is studied considering a single-user and multi-user FSO system with OOK modulation in an additive white Gaussian noise (AWGN) channel. Other applications of DNN in FSO involve channel estimation as demonstrated in Ref [35], imperfect CSI detection [36], blind CSI detection [37] as well as perfect CSI detection [38]. DNN is used in Ref [39] to detect the presence of an eavesdropper and in Ref [40] to decode polar codes, considering the FSO-SISO system with OOK.

1.2 | Practical aspects

Conventional communication systems are composed of multiple individual processing blocks, each performing an isolated task, for example, constellation shaping, channel estimation, and detection. However, isolated processing does not achieve the best possible end-to-end complexity-performance trade-off [33, 34]. Although joint signal processing achieves gains, it leads to computationally complex systems. End-to-end deep learning (DL) optimises the communication system design to achieve efficient performance without requiring a rigid structure [33–35].

The main reason that DL may guide the design of practical FSO communication systems is that it achieves good performance due to its lower complexity and higher processing speed. Nowadays, there are many electronic boards in different platforms (even in mobile platforms) that can be used for deploying real time processing of the proposed DL-based structures. For instance, GPUs and NPUs are available, which can accelerate the processing of the proposed DL-based structures. Let us consider the deployment of high-order

T A B L E 1 Comparison between different works on machine/deep learning for the FSO system

Category	Ref	Application	Transmitter	Receiver	System model	Channel model	Complexity analysis
Machine learning-based algorithms	[16]	Detector to combat noise and scintillation	OOK	SVM	SISO	Log-normal	No
	[17]	Optical power estimator	OOK	Decision tree	SISO	Experimental	No
	[18]	Correcting distortion in a sensor-less adaptive optic system	Phase-only spatial light modulator	ANN	SISO	Experimental	No
	[19]	Mitigating the atmospheric turbulence effects at the beam level without the need for an adaptive optic system	Phase-only spatial light modulator	ANN	SISO	Experimental	No
CNN-based algorithms	[20]	Detector	OOK	CNN	SISO	Gamma–Gamma	No
	[21]	Detector	OOK	CNN	SISO	Gamma–Gamma	No
	[22]	Detector	OOK	CNN	SISO	Gamma–Gamma	No
	[23]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[24]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[25]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[26]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[27]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[28]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Experimental	No
	[29]	Demodulator for OAM-SK	OAM-SK	3 dimensional CNN	SISO	Gamma–Gamma	No
	[30]	Demodulator for OAM-SK	OAM-SK	CNN	SISO	Pointing error	No
	[31]	Demodulator for OAM-SK	OAM-SK	Multi-CNN	SISO	Gamma–Gamma	No
	[32]	Reducing the atmospheric turbulence-induced power loss	Phase-only spatial light modulator	AlexNet-CNN	SISO	Experimental	No
DNN-based algorithms	[33]	Autoencoder	DNN	DNN	SISO	Log-normal	No
	[34]	Autoencoder	DNN	DNN	SISO	Gaussian	No
	[35]	Channel estimation	M-QAM	DNN	SISO	Gamma–Gamma	Yes
	[36]	Detector for imperfect CSI	M-QAM	DNN	SISO	Gamma–Gamma	No
	[37]	Detector for blind CSI	OOK	DNN	SISO	Log-normal/Gamma–Gamma/Negative exponential	Yes
	[38]	Detector for perfect CSI	OOK	DNN	SISO	Gamma–Gamma	Yes
	[39]	Eavesdropper detection	OOK	DNN	SISO	Experimental	No
	[40]	Decoder of polar codes	OOK	DNN	SISO	Gamma–Gamma	No
	This work	Joint (and individual) constellation shaping-detection	DNN	DCNN	MIMO	Gamma–Gamma	Yes

Abbreviations: ANN, artificial neural network; CNN, convolutional neural network; CSI, channel state information; DNN, deep neural network; FSO, free space optical; MIMO, multi-input multi-output; M-QAM, M-ary quadrature amplitude modulation; OAM, orbital angular momentum; OOK, on-off keying; SISO, single-input single-output; SK, shift keying; SVM, support vector machine.

modulation formats in FSO systems, which is presently seen as a promising way of increasing spectral efficiency. Several conventional implementation options for detection are possible, among which ML is the best. However, ML's complexity increases with the modulation order; catering to this interest, DL provides a fast way of detecting optical signals with high-order modulation.

1.3 | Contributions and novelties

The majority of the above-mentioned studies, including those on CNN, did not explore complexity, and their provided algorithms are inefficient in terms of computation. One of the main missions of this study is to propose low complexity DL-based transmitter, receiver, and transceiver structures. We present

low complexity transmitter, receiver, and transceiver structures and compare their performance complexity with each other. Moreover, the proposed methods are compared with conventional methods in terms of performance and complexity. Furthermore, this study presents a comprehensive investigation of transmitter, receiver, and transceiver learning in FSO systems. Note that the aim of using DL at the transmitter, receiver, and transceiver sides is constellation shaping, detection and joint constellation shaping and detection, respectively.

The existing DL-based structures considered the FSO-SISO system. In this study, we investigate both SISO and MIMO FSO systems. To have a comprehensive investigation, different combining techniques such as maximum ratio combining (MRC), equal gain combining (EGC), and selective combining (SC) have been considered. In addition, a wide range of atmospheric turbulence regimes from weak to strong with different modulation orders have been studied in this work. Therefore, the main contributions and novelties of this work are summarised as follows:

- Designing low complexity DL-based transmitter, receiver, and transceiver structures for the FSO system.
- Providing complexity analysis for DL applications in the FSO system.
- Considering transmitter learning in the FSO system.
- Presenting a comprehensive investigation over DL applications in the FSO system.
- Investigating DL applications in the FSO-MIMO system (and deploying, MRC, EGC, and SC schemes).
- Considering a wide range of atmospheric turbulences from weak to strong and different modulation orders.

The rest of this work is organised as follows: Section 2 describes channel and system models. Section 3 describes the system model of the proposed structures. Section 4 presents simulation results, and Section 5 is the conclusion of this work.

2 | FSO CHANNEL AND SYSTEM MODEL

2.1 | FSO channel model

The most important FSO channel effects include weather conditions, pointing error, atmospheric turbulence, and path loss. Considering a short-range link in a clear atmosphere, the major challenges faced by the FSO links are atmospheric turbulence and pointing error. Note that the pointing error can be removed by properly fastening the FSO transceiver. Various statistical distributions have been used in the literature to model the atmospheric turbulence effects [41–49]; among them, Gamma–Gamma distribution has the highest accompany with the actual results for weak to strong atmospheric turbulence regimes [2, 47, 50, 51]. In this study, the atmospheric turbulence effect is modelled by Gamma–Gamma distribution, which has the following probability distribution function:

$$f(I) = \frac{2(\alpha\beta)^{\frac{\alpha+\beta}{2}}}{\Gamma(\alpha)\Gamma(\beta)} I^{\frac{\alpha+\beta}{2}-1} K_{\alpha-\beta}\left(2\sqrt{\alpha\beta I}\right); \ I > 0, \quad (1)$$

where I is the atmospheric turbulence intensity, $\Gamma(.)$ is a well-known gamma function, $K.(.)$ is a modified Bessel function of the second kind and $1/\beta$ and $1/\alpha$ are the variances of the small and large scale eddies, respectively. Moreover, $\alpha = \left[\exp(0.49\sigma_R^2/(1+1.11\sigma_R^{12/5})^{7/6}) - 1\right]^{-1}$, $\beta = \left[\exp(0.51\sigma_R^2/(1+0.69\sigma_R^{12/5})^{5/6}) - 1\right]^{-1}$, where $\sigma_R^2 = 1.23C_n^2 k^{7/6} z^{11/6}$ is the Rytov variance, $k = 2\pi/\lambda$ is the optical wavenumber, and z is the propagation distance [52, 53].

2.2 | FSO system model

As depicted in Figure 1, in the considered FSO system model, the same FSO signal is simultaneously transmitted from N_t optical transmitters, received by N_r optical detectors, and combined by MRC, EGC, or SC. We assume coherent detection at the receiver side; therefore, the phase of the received signal can be detected. Moreover, equal power is allocated to the transmitters, that is, $P_i = P/N_t$, $i = 1, ..., N_t$. Considering x, as the transmitted FSO signal, the received signal at the ith; $i = 1, ..., N_r$ receive aperture can be expressed as

$$y_i = R \sum_{j=1}^{N_t} I_{i,j} x + n_i, \quad (2)$$

where n_i is AWGN at the input of ith receive aperture, with zero mean and variance σ^2; $I_{i,j}$ is the atmospheric turbulence intensity of the link between the jth transmitter and the ith receive aperture. Moreover, $R = \eta q/hf$ is the photo detector responsibility, where η is the quantum efficiency of the photodetector, q is the electron's charge, h is Planck's constant, and f is the optical frequency. We consider receivers with limited background noise in which the shot noise created by background radiation is dominant compared to other noise components such as thermal noise, dark noise, and signal-dependent shot noise. Therefore, the noise term is modelled as signal-independent AWGN [54].

There are several copies of the transmitted signal available at the receiver, which are combined to improve the system capacity and coverage area [55]. Various combining schemes are available; among them MRC, EGC, and SC are some of the well-known combiners [53]. In SC, only the received apertures are selected so that its received signal has the maximum signal-to-noise ratio. Assuming p as the index of the selected receive aperture by SC, the ML receiver for SC becomes

$$\hat{x}_u = \underset{\tilde{x}_u}{\operatorname{argmin}} \left| y - R \sum_{j=1}^{N_t} I_{p,j} \tilde{x}_u \right|^2, \quad (3)$$

where $\tilde{x}_u$ is a symbol of the transmitted constellation map. In EGC, all the received signals are co-phased and added together without any weighting. Since the FSO channel only takes real
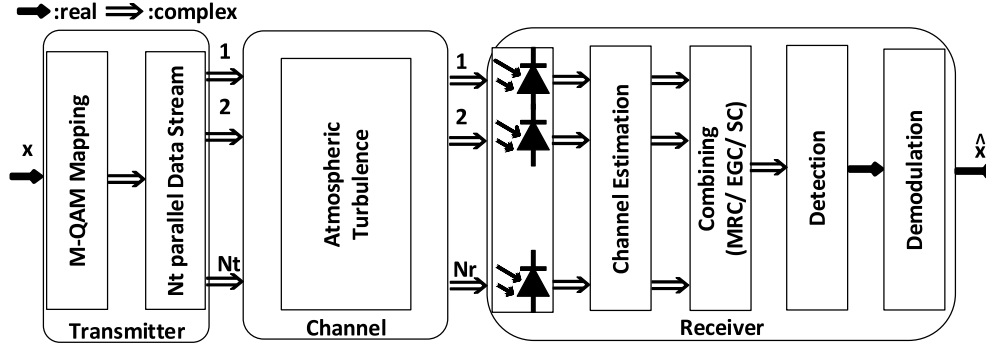
FIGURE 1 Conventional free space optical Communication system. EGC, equal gain combining; MRC, maximum ratio combining; SC, selective combining

values, EGC can be performed by adding the received signals and can be expressed as follows [56]:

$$y = \sum_{i=1}^{M} y_i = \frac{R}{N_t} \sum_{i=1}^{M} \sum_{j=1}^{N_t} I_{i,j} x_u + \sum_{i=1}^{M} n_i. \quad (4)$$

After combing the received signals with the EGC method, the ML receiver for the EGC will be as follows [56]:

$$\hat{x}_u = \operatorname*{argmin}_{\tilde{x}_u} \left| y - \frac{R}{N_t} \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} I_{i,j} \tilde{x}_u \right|^2 \quad (5)$$

In MRC, all of the received signals are rotated and weighted according to the phase and strength of the channel such that all of the received signals are combined to yield the maximum signal to noise ratio. The FSO channel only takes real values; therefore, MRC can be performed by adding the weighted received signals, and can be expressed as follows [56]:

$$y = \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} I_{i,j} y_i = R^2 \sum_{i=1}^{N_r} \left(\sum_{j=1}^{N_t} I_{i,j} \right)^2 x + R \sum_{i=1}^{N_r} \sum_{j=1}^{N_t} I_{i,j} n_i. \quad (6)$$

After combing the received signals with the MRC method, the ML receiver for the MRC is given by [56]

$$\hat{x} = \operatorname*{argmin}_{\tilde{x}} \left| y - R^2 \sum_{i=1}^{N_r} \left(\sum_{j=1}^{N_t} I_{i,j} \right)^2 \tilde{x} \right|^2. \quad (7)$$

3 | PROPOSED DL-BASED FSO SYSTEMS

Deep learning is an evolution of machine learning, which uses a programmable neural network that enables machines to make accurate decisions without help from humans. This technique

is successfully applied in different OC applications [57]. Deep learning is an established technique for learning complex relationships between the received signal and impairments. Among DL algorithms, the DNN is the most widely used algorithm in OC applications, which provides favourable performance and can be considered as an alternative for conventional methods [57]. The deep neural network learns the relationship between the input data and target output by using several hidden layers, each consisting of multiple connected neurons by some weights, biases and activation functions that represent the importance of each connection. The dimensionality of DNN can be reduced by combining it with CNN. This is called network in the DL community and is widely investigated in image classification [38].

In this study, three DL-based FSO systems are proposed in which DL is deployed at the receiver (classical transmitter/DL-based detector), transmitter (DL-based transmitter/ML detector), and transceiver sides (DL-based transmitter/DL-based detector) for detecting, constellation shaping, and joint constellation shaping-detection, respectively. In the following sections, these proposed structures are described in detail.

3.1 | Classical transmitter/DL-based detector

The proposed classical transmitter/DL-based detector system model is presented in Figure 2a. Considering x as the transmitted M-ary symbol, it is first converted to the one-hot vector $s \in S = \{e_i | i = 1, \dots, M\}$, where e_i equals 1 at row i and else 0, then mapped on an M-ary quadrature amplitude modulation (M-QAM) constellation. The one-hot vector is composed since a target is required for training the DL algorithm; the aim of the training is to reduce the difference between the DL algorithm output (the detected symbol) and the target (the transmitted symbol). Then, N_t copies of the mapped symbol are transmitted simultaneously from N_t FSO transmit apertures. The transmitted signal is encountered by the Gamma–Gamma atmospheric turbulence channel. In the receiver side, the received optical signal is covenanted to the electrical signal in each photodetector. Different received copies of the transmitted signals are then combined at the receiver side using MRC, EGC or SC methods. Note that the
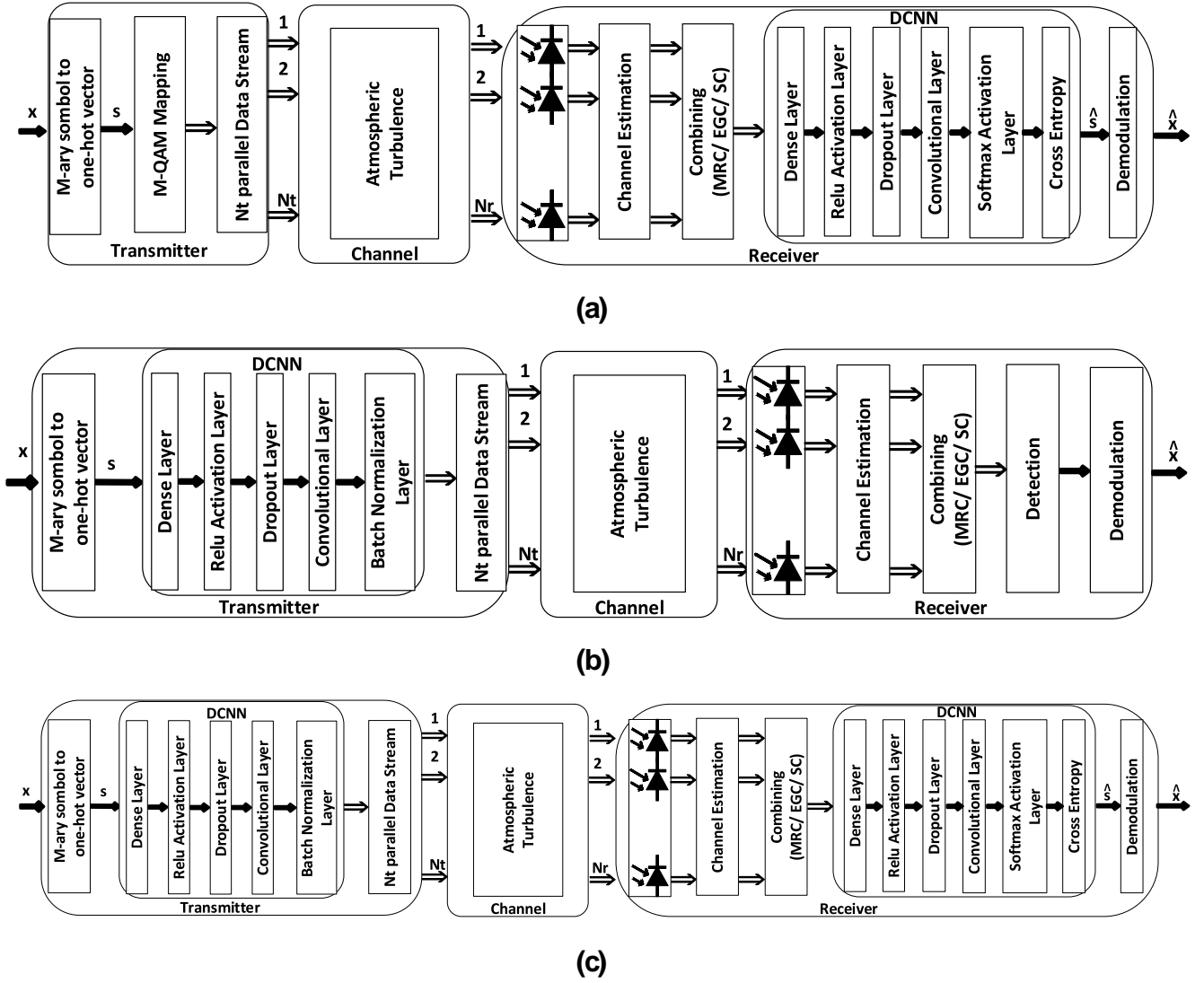
FIGURE 2 The proposed (a) Classical transmitter/deep learning (DL)-based detector; (b) DL-based transmitter/maximum likelihood detector; and (c) DL-based transmitter/DL based detector structures. DCNN, dense convolutional neural network; EGC, equal gain combining; MRC, maximum ratio combining; SC, selective combining

aforementioned combing methods need the CSI at the receiver side. After combining the received signal, it is entered in the detector. Our detector is inspired from Ref [38] where a convolutional layer is used in DNN for the sake of dimensionality reduction and speedup. To distinguish it from DNN and CNN, we call it the dense CNN (DCNN). The dense layer does not support complex numbers; therefore, the real and imagery parts of the combined signal should be separately fed. The dense layer is composed of two input neurons and N_{hid} hidden layers each with N_{neu} hidden neurons. The output of each layer is multiplied by a weight matrix (W), summed by a bias vector (b), and passed through an activation function ($\alpha(.)$). Note that after each hidden layer, a drop out layer is used for dimensionality reduction. In the output layer of the DCNN, M 1×1 convolutional filters are used. Therefore, the DCNN is different from CNN wherein the convolutional layers are prior to the dense layers.

The weights and biases should be adjusted such that the difference between the output and the target is minimised. In order to adjust these parameters, the DCNN algorithm should be trained. The first step is selecting and tuning its hyperparameters [57]. The DCNN hyperparameters include the sample size to batch size ratio, layer type, number of layers, number of neurons, activation function, loss function, optimiser, learning rate, and number of iterations. The sample size to batch size ratio is important because entering the whole data at once into the DCNN leads to underfitting, while dividing it into several batches helps the DCNN to better understand the data structure. The number of layers, as well as neurons, should be adjusted by trial and error, and there is no specific rule for tuning them. The activation functions are extended during the time and according to complexity, accuracy, and timing demands, there is a trade-off between them; however some of them such as tanh, sigmoid, relu are shown to be proper for DL in OC applications.

After selecting and tuning the structural hyperparameters, it is necessary to define the hyperparameters directly related to the training aspect. The inputs of each layer of the DCNN are multiplied by the corresponding weights, added by biases, summed, and then passed through the activation function. The outputs of each layer are the inputs of the next layer, and the procedure continues until it reaches the last layer. Considering the original one-hot vector at the transmitter as s and the output vector of the DCNN as $\hat{s}$, the aim is to reduce the difference between s and $\hat{s}$. Therefore, a loss function should be defined and calculated for each individual transmitted symbol and expected over the whole batch size. The loss function can be defined as in Ref [57]

$$L(\boldsymbol{\theta}) = \frac{1}{K} \sum_{k=1}^{K} \left[l^{(k)}(\boldsymbol{s}, \hat{\boldsymbol{s}}) \right], \tag{8}$$

where $\boldsymbol{\theta}$ is a vector including weights and biases, K is the batch size, and $l(.,.)$ is the loss function. The cross-entropy loss function is considered in this study and is defined as in Ref [57]

$$l(\boldsymbol{s}, \hat{\boldsymbol{s}}) = -\sum_i s_i \, log(\hat{s}_i). \tag{9}$$

Several algorithms are available for finding $\boldsymbol{\theta}$ that minimise the loss function. One of the most popular algorithms is the stochastic gradient descent (SGD) that obtains the $\boldsymbol{\theta}$ iteratively by the following formulation [58]:

$$\boldsymbol{\theta}^{(m+1)} = \boldsymbol{\theta}^{(m)} - \eta \nabla_{\boldsymbol{\theta}} \tilde{L} \left(\boldsymbol{\theta}^{(m)} \right) \tag{10}$$

where $\eta > 0$ is the learning rate, m is the iteration number, and $\nabla_{\boldsymbol{\theta}} \tilde{L}(.)$ is the estimate of the gradient. The error derivation $(\nabla_{\boldsymbol{\theta}} \tilde{L}(\boldsymbol{\theta}^{(j)}))$ is fed back to the DCNN as an updating guide. The SGD includes many extensions; among them, the Adam algorithm is the state-of-the-art algorithm with enhanced convergence [57]. The Adam algorithm is used for optimization during the training process in this work.

3.2 | DL-based transmitter/ML detector

The proposed DL-based transmitter/ML detector system model is depicted in Figure 2b. The idea of using DL on the transmitter is to learn a waveform representation that is robust to the FSO channel impairments. The trained transmitter can be viewed as a look-up table that simply maps the input message to one of the optimised blocks. The generated M-ary symbol x is first converted to the one-hot vector s, then entered into a DCNN. The dense layer is composed of M input neurons and N_{hid} hidden layers each with N_{neu} hidden neurons. After each hidden layer, a drop out layer is used. At the output layer of the DCNN, 2 1×1 convolutional filters are used. Complex summation of the neurons at the output layer results in a complex number, which stands for the location of symbol x in the constellation. Note that this value is limited to

transmitter power constraint; therefore, a batch normalisation layer is added after the convolutional layer. The DCNN maps the one-hot vector s onto its location on the constellation. Then, N_t copies of the mapped symbol are transmitted simultaneously through N_t transmit apertures, encountered by Gamma–Gamma atmospheric turbulence, added by AWGN with zero mean and σ^2 variance, and combined by MRC, EGC or SC at the receiver side. The combined signal is then passed through an ML detector. The detected symbol $\hat{x}$ is converted into the one-hot vector $\hat{s}$. In order to minimise the difference between the transmitted one-hot vector (s) and the detected one-hot vector ($\hat{s}$), the weights and biases of the DCNN should be adjusted by training. The first step towards this end is selecting and tuning the hyperparameters. Then, a loss function should be defined based on s, $\hat{s}$, and tuned hyperparameters as expressed in Equation (8). The last step is iteratively minimising the loss function using the Adam algorithm as expressed in Equation (10).

3.3 | DL-based transmitter/DL-based detector

The proposed DL-based transmitter/DL-based detector system model is presented in Figure 2c. Considering x as the generated M-ary symbol, it is first converted into a one-hot vector then entered into a DCNN with the same structure as explained in Section 2.2. The complex summation of the output neurons stands for the location of the mapped symbol in the constellation. Then, N_t copies of the mapped symbol are transmitted simultaneously through N_t transmit apertures, encountered by Gamma–Gamma atmospheric turbulence, added by AWGN with zero mean and σ^2 variance, and combined by MRC, EGC or SC at the receiver side. The combined signal is entered into a DCNN in the same way as explained in Section 2.1. This end-to-end DL structure is used for joint constellation-shaping detection to reduce the effect of atmospheric turbulence. Considering the transmitted one-hot vector (s) and the detected one-hot vector ($\hat{s}$), the weights and biases of the transceiver DCNN should be adjusted by training. Therefore, first, the hyperparameters should be selected and tuned. Then, based on s, $\hat{s}$, and tuned hyperparameters, a loss function should be defined as shown in Equation (8). Finally, the loss function should be iteratively minimised using the Adam algorithm as expressed in Equation (10).

4 | RESULTS AND DISCUSSIONS

The training of the proposed DL algorithms is performed in a Python/Tensorflow environment. A batch of random M-ary symbols is created and used for training the DL-based structures, as described in Section 3. The DL-based hyperparameters are tuned manually, and the values of the tuned hyperparameters and other simulation parameters are given in Table 2. In this section, the simulation results of testing the proposed DL-based structures are presented.

4.1 | Complexity analysis

Table 3 presents the computational complexity of the DL-based transmitter and ML detector. Note that the complexity of the classical transmitter and DL-based transmitter are equal, as they both use lookup tables of the same size. The training of DL algorithms is carried out once off-line; therefore, it is not required to calculate the computational complexity of the training phase. In addition, it is sufficient to take into account the forward propagation; the backward propagation is not important, as it happens only in the training phase and not in the test phase. The complexity of the DL-based detector is equal to the number of multiplications and summations of the DCNN. The DL-based detector is composed of an input layer with two neurons, N_{hid} hidden layers each with N_{neu} hidden neurons, and an output layer with M 1×1 kernels. Therefore, its complexity becomes $(2 \times N_{neu} + N_{hid} \times N_{neu} \times N_{neu} + M \times N_{neu})/2$ multiplications and $(N_{neu} \times 3 + N_{hid} \times N_{neu} \times (N_{neu} + 1) + \times N_{neu})/2$ summations, where M is the modulation order. The complexity of the ML detector is $7 \times M - 3$ multiplications and $4 \times M - 1$ summations.

In order to have a better understanding about the computational complexity of conventional and the proposed DL-based FSO systems, Figure 3 displays the run time of the classical transmitter/ML detector, classical transmitter/DL-based detector, DL-based transmitter/ML detector, and DL-based transmitter/DL-based detector structures as functions of the modulation order, for the SISO scheme with a strong atmospheric turbulence regime when Es/N0 = 10 dB. The DL-based detector has lower complexity than the ML detector, and the complexity difference between the DL-based detector and ML detector increases by the modulation order. As mentioned in Table 3, the complexity order of the DL-based detector and ML detector are $O(M/2)$ and $O(M)$, respectively.

The complexity of the classical transmitter and DL-based transmitter is almost the same. The classical transmitter searches in the lookup table for constellation shaping, and the DL-based transmitter does some multiplications and additions on the input one-hot vector and acts like a lookup table. To be more specific, at high modulation orders, the classical transmitter has a bit lower complexity than the DL-based transmitter (because in this range the size of the input one-hot vector is high and increases the complexity); however, at low modulation orders it is inverse (because at this range the size of the input one-hot vector is low).

Note that the complexity difference of conventional and DL-based FSO systems is mainly on the receiver side. For instance, considering the modulation order $M = 16$ where the classical transmitter and DL-based transmitter have the same complexity, the DL-based detector is almost 2 times faster than the ML detector. This difference increases when the modulation order is increased; for example, in $M = 64$ and $M = 256$, the DL-based detector is 3 times and 7.5 times faster than the ML detector, respectively, when using the classical transmitter and 2.8 times and 6.3 times faster than the ML detector when using the DL-based transmitter.

4.2 | Performance analysis

To have a better conclusion, we investigate the performance of the conventional and DL-based FSO systems in this part. Note that the ML detector, which we used as a reference conventional detector, is the optimum detector. In addition, the M-ary QAM, due to its good performance, is an outstanding modulation format in most of the communication standards.

In Figure 4, the symbol error rates (SER) of the classical transmitter/ML detector, classical transmitter/DL-based detector, DL-based transmitter/ML detector, and the DL-based transmitter/DL-based detector structures are plotted as functions of Es/N0, for SISO and MIMO (MRC 2×2) schemes, for a strong atmospheric turbulence regime. Considering the computation of SER, it should be noted that the output one-hot vector of the DL algorithm, $\hat{s}$, is converted to an M-ary symbol, $\hat{x}$; the SER is the difference between the transmitted and received M-ary symbols, x and $\hat{x}$. As seen in Figure 4, the proposed DL-based structures achieve the performance of the state-of-the-art classical transmitter/ML detector. Note that perfect channel estimation is carried out at the receiver side of these structures. In this situation, the DL-based detector and the ML detector both minimise the difference between transmitted and received symbols and achieve the same performance. In addition, the DL-based transmitter achieves the same performance as the classical transmitter, since the atmospheric turbulence effect is removed at the receiver. It might be interesting to consider that the obtained results are achieved by simple hyperparameter tuning and with the use of a simple structure. According to these results, and considering the fact that hyperparameter tuning is performed manually, it can be concluded that the proposed DL-based structures work perfectly and efficiently.

Figure 5 plots the SER of the classical transmitter/ML detector and DL-based transmitter/DL-based detector structures as functions of Es/N0, for SISO and MIMO (MRC 2×2) schemes, for different atmospheric turbulence regimes. As seen in Figure 5, the DL-based transmitter/DL-based detector is a bit better than the classical transmitter/ML detector. Figure 4, shows that the DL-based detector has the same performance as the ML detector. Accordingly, the fact that the DL-based transmitter/DL-based detector is a bit better than the classical transmitter/ML detector shows that the DL-based transmitter can find a constellation shape that performs a bit better than M-QAM. Note that M-QAM is an outstanding modulation being used in many communication standards. The hyperparameter tuning is performed based on the SISO structure, manually and based on previous knowledge. Despite this fact and the complexity of MIMO and the 16-ary constellation, the proposed DL-based transmitter/DL-based detector achieves similar performance at all atmospheric turbulence regimes in the MIMO scheme. It shows the flexibility of the proposed DL-based transmitter/DL-based detector, which means it is not required to consider all scenarios for tuning and the tuning can be carried out considering a desirable scenario; the DCNN will

T A B L E   2   The parameters used in simulations and their values

Parameter	Value
Modulation order	16
Number of hidden layers (N_{hid})	2
Number of hidden neurons (N_{neu})	10
Batch size	2^{16}
Sample size to batch size ratio	16
Number of iterations (epochs)	300
Activation function	Relu/Softmax
Layer type	Dense/Convolutional
Loss function	Cross entropy
Optimiser	Adam
Learning rate	0.005
Gamma–Gamma atmospheric turbulence intensity	Strong ($\alpha = 4.2, \beta = 1.4$)/Moderate ($\alpha = 4, \beta = 1.9$)/Weak ($\alpha = 11.6, \beta = 10.1$)
Photo detector responsibility	$R = 1$

T A B L E   3   Computational complexity of the DL-based detector and the ML detector

Method	Number of summations and diffractions	Number of multiplications and divisions	Complexity order
DL based detector	$\frac{(N_{neu} \times 3 + N_{hid} \times N_{neu} \times (N_{neu}+1) + M \times N_{neu})}{2}$	$\frac{2 \times N_{neu} + N_{hid} \times N_{neu} \times N_{neu} + M \times N_{neu}}{2}$	$O(M/2)$
ML detector	$7 \times M - 3$	$4 \times M - 1$	$O(M)$

Abbreviations: DL, deep learning; ML, maximum likelihood.
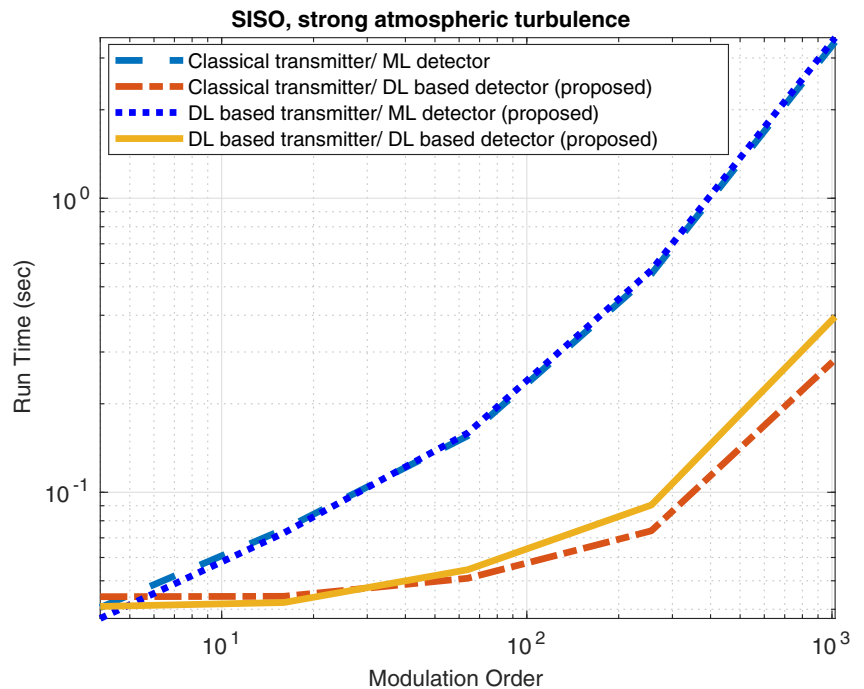
F I G U R E   3   Run time of the classical transmitter/maximum likelihood (ML) detector, classical transmitter/deep learning (DL) based detector, DL based transmitter/ML detector, and DL based transmitter/DL based detector structures as functions of the modulation order, for the single-input single-output (SISO) scheme, and for a strong atmospheric turbulence regime, when Es/N0 = 10 dB
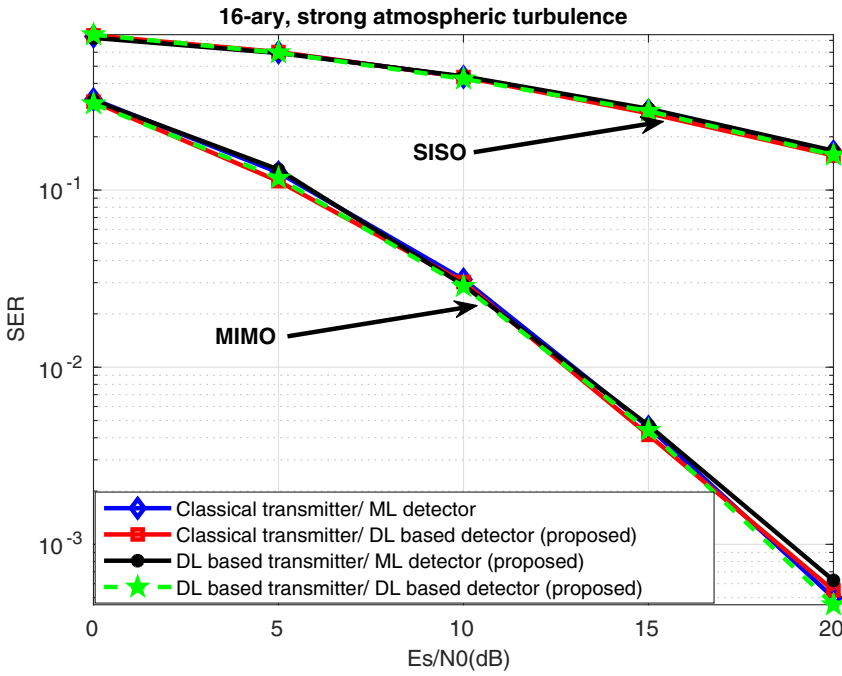
FIGURE 4 Symbol error rates (SERs) of the classical transmitter/maximum likelihood (ML) detector, classical transmitter/deep learning (DL) based detector, DL based transmitter/ML detector, and the DL based transmitter/DL based detector structures as functions of Es/N0, for single-input single-output (SISO) and multi-input multi-output (MIMO) (maximum ratio combining 2×2) schemes, for a strong atmospheric turbulence regime
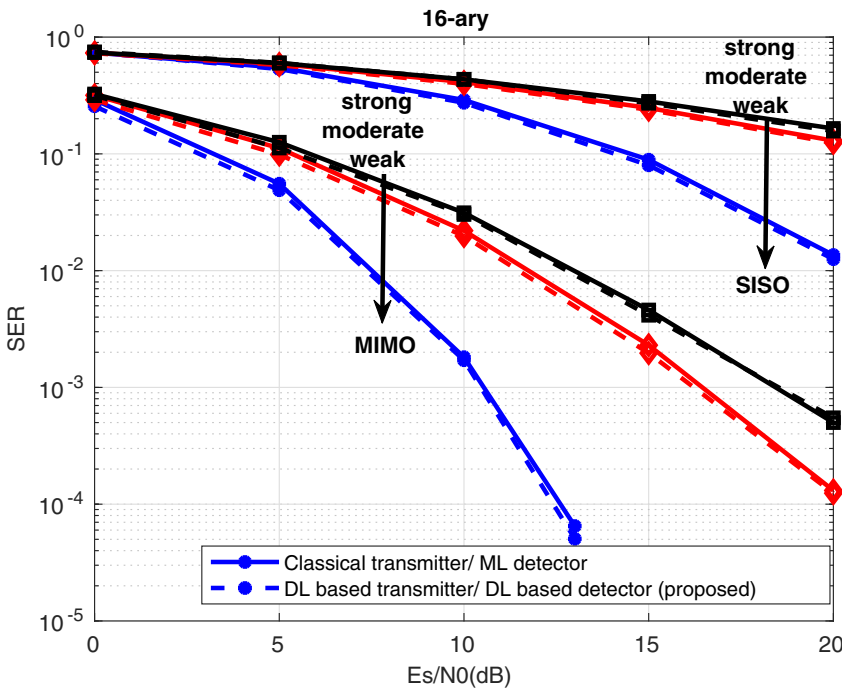

FIGURE 5 Symbol error rates (SERs) of the classical transmitter/maximum likelihood (ML) detector and deep learning (DL) based transmitter/ DL based detector structures as functions of Es/ N0, for single-input single-output (SISO) and multi-input multi-output (MIMO) (MRC 2×2) schemes, for different atmospheric turbulence regimes

work well in all scenarios. This shows that this structure is robust to any changes in the structure of the system, and by a simple training procedure it can be adapted to any new structure. Therefore, the DL-based transmitter/DL-based detector, due to its much lower complexity, is a good choice for working in situations with varying channel conditions.

Figure 6 shows the SERs of the classical transmitter/ML detector and DL-based transmitter/DL-based detector structures as functions of Es/N0, for MIMO (2×2) with different

combining schemes, for a strong atmospheric turbulence regime. As seen in Figure 6, in both the classical transmitter/ ML detector and DL-based transmitter/DL-based detector structures, MRC outperforms EGC, and EGC outperforms SC. In addition, the DL-based transmitter/DL-based detector performs the same as the classical transmitter/ML detector structure for different combining schemes. However, it should be considered that the number of transceiver apertures is only 2, and the investigation is carried out over a strong atmospheric turbulence regime; therefore, MRC, EGC, and SC are
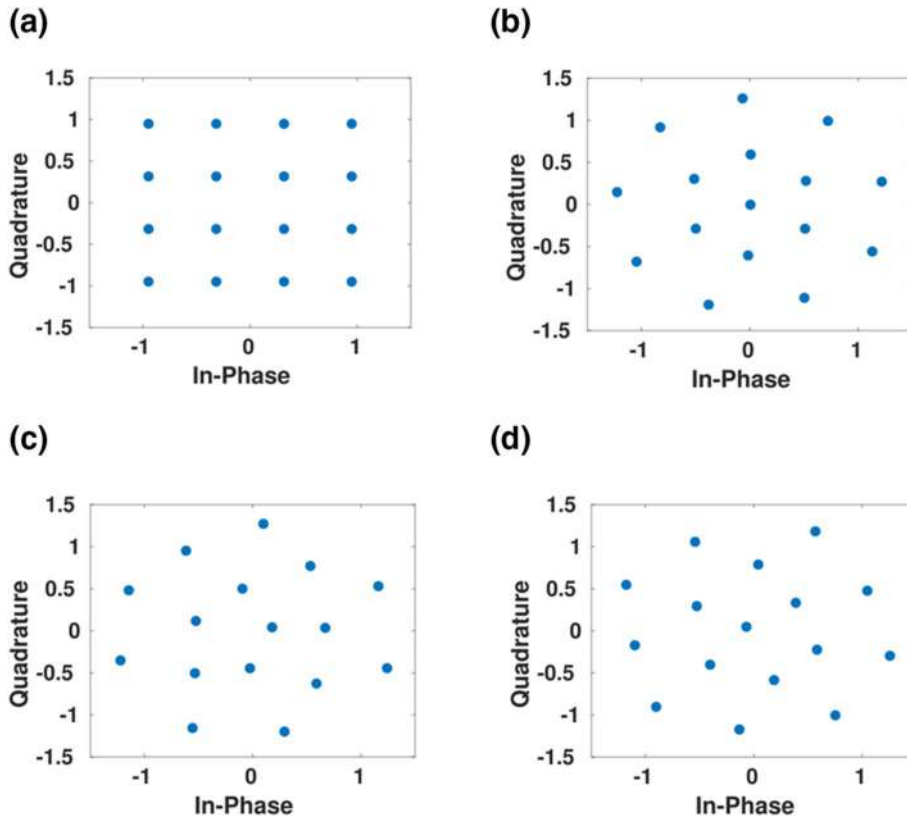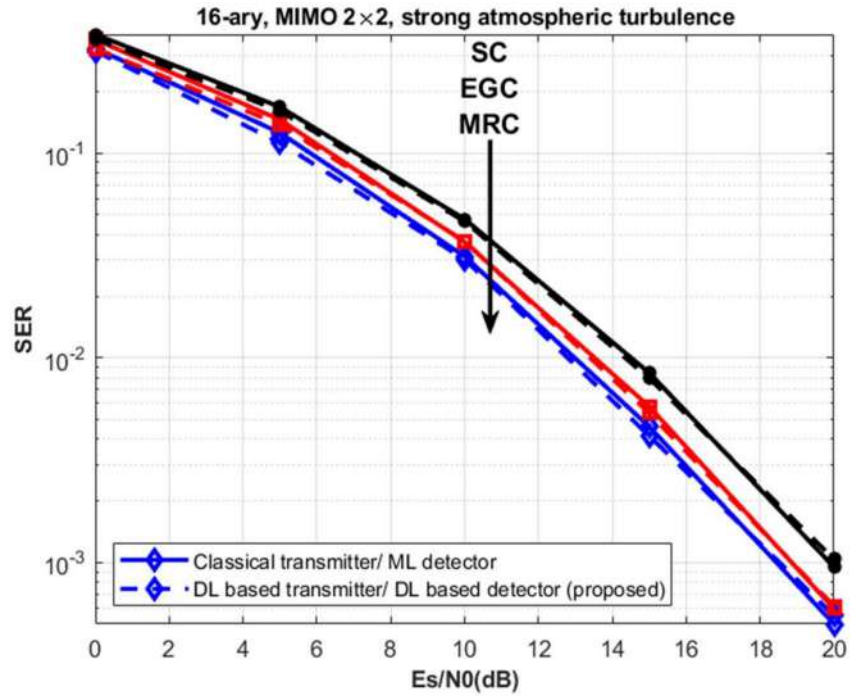
FIGURE 7 The 16-ary constellation of transmitted signal for (a) QAM transmitter, and the proposed deep learning based transmitter in (b) Weak; (c) Moderate; and (d) Strong atmospheric turbulence, with Es/N0 = 10 dB

not very different from each other. However, an increasing number of apertures would increase this difference. As seen in Figure 6, at low Es/N0, MRC, EGC, and SC have almost the same performance; however, at higher Es/N0, there is almost a fixed difference between them; this is due to the dominancy of noise or Es/N0 in each case.

Figure 7 shows the 16-ary constellation of transmitted signal for a. QAM transmitter and the proposed DL-based transmitter in b. Weak, c. Moderate, and d. Strong atmospheric turbulence, with Es/N0 = 10 dB. In a perfect CSI scenario, the receiver refines the received signal constellation and then detects the received signal. The constellation shaper considers the channel as an AWGN (with variance of $\sigma^2/|I|^2$). In the AWGN channel, the optimum constellation is a distribution of points with equal distances. This criterion can be seen in M-QAM as well as in Figure 7, and that is why the performance of the classical transmitter and DL-based transmitter is the same.

5 | CONCLUSION

In this study, three DL-based structures were proposed for mitigating the atmospheric turbulence in FSO communication systems. In the proposed structures, DL was used at the transmitter, receiver, and transceiver sides, respectively, for constellation shaping, detection, and joint constellation-shaping detection. Considering the SISO/MIMO structure and a wide range of atmospheric turbulences, from the weak to the strong regime, the performance of the proposed structures was compared with that of the ML detector. Simulation results showed that the proposed DL-based methods achieved the ML performance with lower complexity. The DL-based transmitter/ML detector had almost the same complexity as the ML detector, while the DL-based detector was almost 2, 3, and 7.5 times faster than the ML detector in modulation orders of 16, 64, and 256, respectively.

CONFLICT OF INTEREST
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this study.

DATA AVAILABILITY STATEMENT
There was no data for this work.

ORCID
Mohammad Ali Amirabadi https://orcid.org/0000-0003-4190-3380

REFERENCES
1. Majumdar, A.K.: Advanced Free Space Optics (FSO): A Systems Approach, vol. 186. Springer (2014)
2. Bhatnagar, M.R., Ghassemlooy, Z.: Performance evaluation of FSO MIMO links in Gamma–Gamma fading with pointing errors. In: IEEE International Conference on Communications (ICC), pp. 5084–5090 (2015)
3. Jaiswal, A., Bhatnagar, M.R.: Free-space optical communication: a diversity-multiplexing tradeoff perspective. IEEE Trans. Inf. Theor. 65(2), 1113–1125 (2018)
4. Arti, M.K., Jain, A.: A simple approximation of FSO link distribution and its applications. In: 2020 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), pp. 1–3. Dec 2020
5. Arti, M.K., Bhatnagar, M.R.: Maximal ratio transmission in AF MIMO relay systems over Nakagami-*m* fading channels, IEEE Trans. Veh. Technol. 64(5), 1895–1903 (2014)
6. Arti, M.K., Bhatnagar, M.R.: Imperfect CSI based two-way AF MIMO relaying of OSTBC. In: 2014 IEEE International Conference on Communications (ICC), pp. 4430–4435. June 2014
7. Amran, N.A., et al.: Deep learning based signal detection for OFDM VLC systems. In: IEEE International Conference on Communications Workshops (ICC Workshops), pp. 1–6 (2020)
8. Khan, F.N., et al.: An optical communication's perspective on machine learning and its applications. J. Lightwave Technol. 37(2), 493–516 (2019)
9. Yuan, Y., et al.: SVM detection for superposed pulse amplitude modulation in visible light communications. In: 10th International Symposium on Communication Systems, Networks and Digital Signal Processing (CSNDSP), pp. 1–5 (2016)
10. Pashazanoosi, M., Nezamalhosseini, S.A., Salehi, J.A.: LiFi grid: a machine learning approach to user-centric design, Appl. Opt. 59(28), 8895–8901 (2020)
11. Han, Y., et al.: An SVM-based detection for coherent optical APSK systems with nonlinear phase noise. IEEE Photonics J. 6(5), 1–10 (2014)
12. Thrane, J., et al.: Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals, J. Lightwave Technol. 35(4), 868–875 (2017)
13. Wu, X., Chi, N.: The phase estimation of geometric shaping 8-QAM modulations based on K-means clustering in underwater visible light communication. Opt. Commun. 444, 147–153 (2019)
14. Rottondi, C., et al.: Machine-learning method for quality of transmission prediction of unestablished lightpaths. IEEE/OSA J. Opt. Commun. Netw. 10(2), A286–A297 (2018)
15. Amirabadi, M.A., Kahaei, M.H., Nezamalhosseini, S.A.: Novel suboptimal approaches for hyperparameter tuning of deep neural network [under the shelf of optical communication]. Phys. Commun. 41, 101057 (2020)
16. Zheng, C., Yu, S., Gu, W.: A SVM-based processor for free-space optical communication. In: IEEE 5th International Conference on Electronics Information and Emergency Communication, pp. 30–33 (2015)
17. Tóth, J., et al.: Classification prediction analysis of RSSI parameter in hard switching process for FSO/RF systems. Measurement. 116, 602–610 (2018)
18. Li, Z., Zhao, X.: BP artificial neural network based wave front correction for sensor-less free space optics communication. Opt. Commun. 385, 219–228 (2017)
19. Lohani, S., Glasser, R.T.: Turbulence correction with artificial neural networks. Opt. Lett. 43(11), 2611–2614 (2018)
20. Krenn, M., et al.: Communication with spatially modulated light through turbulent air across Vienna. New J. Phys. 16, 113028 (2014)
21. Ragheb, A., et al.: Identifying structured light modes in a desert environment using machine learning algorithms. Opt. Express. 28(7), 9753–9763 (2020)
22. Lohani, S., Savino, N.J., Glasser, R.T.: Free-space optical ON-OFF keying communications with deep learning. In: Frontiers in Optics, pp. FTh5E-4, Sept 2020
23. Rostami, S., Saad, W., Hong, C.S.: Deep learning with persistent homology for orbital angular momentum (OAM) decoding, IEEE Commun. Lett. 24(1), 117–121 (2019)
24. Hao, Y., et al.: High-accuracy recognition of orbital angular momentum modes propagated in atmospheric turbulences based on deep learning. IEEE Access. 8, 159542–159551 (2020)
25. Liu, J., et al.: Deep learning based atmospheric turbulence compensation for orbital angular momentum beam distortion and communication. Opt. Express. 27(12), 16671–16688 (2018)
26. Tian, Q., et al.: Turbo-coded 16-ary OAM shift keying FSO communication system combining the CNN-based adaptive demodulator. Opt. Express. 26(21), 27849–27864 (2018)

27. Wang, Z., et al.: Efficient recognition of the propagated orbital angular momentum modes in turbulences with the convolutional neural network. IEEE Photonics J. 11, 7903614 (2019)

28. Li, J., et al.: Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication. Opt. Express. 26(8), 10494–10508 (2018)

29. El-Meadawy, S.A., et al.: Performance analysis of 3D video transmission over deep-learning-based multi-coded N-ary orbital angular momentum FSO system. IEEE Access. 9, 110116–110136 (2021)

30. Gong, B., et al.: Recognition of OAM state using CNN based deep learning for OAM shift keying FSO system with pointing error and limited receiving aperture. In: CLEO: QELS_Fundamental Science, pp. JTu3A-169. Optical Society of America, May 2021

31. Li, Z., Su, J., Zhao, X.: Two-step system for image receiving in OAM-SK-FSO link, Opt. Express. 28(21), 30520–30541 (2020)

32. Jin, X., Li, S., Xu, Z.: Compensation of turbulence-induced wavefront aberration with convolutional neural networks for FSO systems. Chin. Opt. Lett. 19(11), 110601 (2021)

33. Zhu, Z.R., et al.: Autoencoder-based transceiver design for OWC systems in log-normal fading channel. IEEE Photonics J. 11(5), 1–12 (2019)

34. Soltani, M., et al.: Autoencoder-based optical wireless communications systems. In: IEEE Globecom Workshops (GC Wkshps), pp. 1–6, Dec 2018

35. Amirabadi, M.A., et al.: Deep learning for channel estimation in FSO communication system. Opt. Commun. 459, 124989 (2020)

36. Amirabadi, M.A., Kahaei, M.H., Nezamalhosseini, S.A.: A deep learning based detector for FSO system considering imperfect CSI scenario. In: 2020 3rd West Asian Symposium on Optical and Millimeter-wave Wireless Communication (WASOWC), pp. 1–5, Nov 2020

37. Amirabadi, M.A., Kahaei, M.H., Nezamalhosseini, S.A.: Deep learning based detection technique for FSO communication systems. Phys. Commun. 43, 101229 (2020)

38. Darwesh, L., Kopeika, N.S.: Deep learning for improving performance of OOK modulation over FSO turbulent channels. IEEE Access. 8, 155275–155284 (2020)

39. Savino, N.J., Lohani, S., Glasser, R.T.: Simulated eavesdropper detection in free-space optics ON-OFF keying with deep learning. In: Frontiers in Optics, pp. FTh5E-6, Sept 2020

40. Fang, J., et al.: Neural network decoder of polar codes with tanh-based modified LLR over FSO turbulence channel. Opt. Express. 28(2), 1679–1689 (2020)

41. Riediger, M.L.B., Schober, R., Lampe, L.: Blind detection of on-off keying for free-space optical communications. In: Canadian Conference on Electrical and Computer Engineering, pp. 001361–001364 (2008)

42. Trinh, P.V., Thang, T.C., Pham, A.T.: Mixed mmWave RF/FSO relaying systems over generalized fading channels with pointing errors. IEEE Photonics J. 9(1), 1–14 (2017)

43. Lopez-Martinez, F.J., Gomez, G., Garrido-Balsells, J.M.: Physical-layer security in free-space optical communications. IEEE Photonics J. 7(2), 1–14 (2015)

44. Zhu, Y.J., et al.: A fast blind detection algorithm for outdoor visible light communications. IEEE Photonics J. 7(6), 1–8 (2015)

45. Anees, S., Bhatnagar, M.R.: Information theoretic analysis of a dual-hop DF based FSO communication system. In: 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), pp. 1–5. May 2016

46. Garg, A., et al.: Performance analysis of erroneous feedback-based MIMO system over Nakagami-m fading channels. IEEE Trans. Commun. 67(8), 5403–5418 (2019)

47. Garg, A., et al.: Imperfect-quantized-feedback-based beamforming for an FSO MISO system over Gamma–Gamma fading with pointing errors. J. Opt. Commun. Netw. 9(11), 1005–1018 (2017)

48. Paul, P., Bhatnagar, M.R., Jaiswal, A.: Alleviation of jamming in free space optical communication over Gamma–Gamma channel with pointing errors. IEEE Photonics J. 11(5), 1–18 (2019)

49. Priyadarshani, R., et al.: Performance of space shift keying over a correlated Gamma–Gamma FSO-MISO channel. In: 2018 11th International Symposium on Communication Systems, Networks & Digital Signal Processing (CSNDSP), pp. 1–6. July 2018

50. Bhatnagar, M.R., Anees, S.: On the performance of Alamouti scheme in Gamma–Gamma fading FSO links with pointing errors, IEEE Wirel. Commun. 4(1), 94–97 (2014)

51. Bhatnagar, M.R.: A one bit feedback based beamforming scheme for FSO MISO system over Gamma–Gamma fading, IEEE Trans. Commun. 63(4), 1306–1318 (2015)

52. Uysal, M., et al.: Optical Wireless Communications: An Emerging Technology. Springer (2016)

53. Khalighi, M.A., Uysal, M.: Survey on free space optical communication: a communication theory perspective. IEEE Commun. Surv. Tutor. 16(4), 2231–2258 (2014)

54. Lapidoth, A., Moser, S.M., Wigger, M.A.: On the capacity of free-space optical intensity channels. IEEE Trans. Inf. Theor. 55(10), 4449–4461 (2009)

55. Bhatnagar, M.R.: Differential decoding of SIM DPSK over FSO MIMO links. IEEE Commun. Lett. 17(1), 79–82 (2012)

56. Bhatnagar, M.R., Ghassemlooy, Z.: Performance analysis of Gamma–Gamma fading FSO MIMO links with pointing errors. J. Lightwave Technol. 34(9), 2158–2169 (2016)

57. Jones, R.T., et al.: Deep learning of geometric constellation shaping including fiber nonlinearities. In: European Conference on Optical Communication (ECOC), pp. 1–3 (2018)

58. O'Shea, T., Hoydis, J.: An introduction to deep learning for the physical layer. IEEE Trans. Cogn. Commun. Netw. 3(4), 563–575 (2017)