Deep Learning for Improving Performance of OOK Modulation Over FSO Turbulent Channels

LAIALY DARWESH, (Member, IEEE), AND NATAN S. KOPEIKA, (Life Senior Member, IEEE)

School of Electrical and Computer Engineering, Ben-Gurion University of the Negev, Beer Sheva 84105, Israel

Corresponding author: Laialy Darwesh (laialy@post.bgu.ac.il)

ABSTRACT Free space optical (FSO) communication technology has become increasingly advanced with capabilities of high speed, high capacity, and low power consumption. However, despite the great potential of FSO, its performance is limited in a turbulent atmosphere. Atmospheric turbulence causes scintillation in the FSO propagated signals, leading to an increase in the bit error rate (BER) performance of the recovered signals at the receiver. In this paper, we demonstrate that the use of deep learning (DL) detection methods could overcome these limitations. We present a new detection method of on-off keying (OOK) modulated signals by using different models of DL over different strength FSO turbulent channels, without the need for prior knowledge of the parameters of the channel. The demonstrated DL decoders improve the performance of the FSO turbulent channel and decrease the power consumption. Moreover, the demonstrated DL models also work faster than maximum likelihood (ML) methods with perfect channel estimation decoders, with even slightly better performance because of the turbulence, thus enabling realization of FSO over turbulent atmospheric channels.

INDEX TERMS Free space optical communication, deep learning, on-off keying modulation, amplitude shift keying modulation, maximum likelihood, channel state information, fully convolutional neural network, fully connected neural network, additive white gaussian noise, intensity modulation, direct detection, photodetector, bit error rate.

I. INTRODUCTION

Free space optical (FSO) communication has gained significant attention in recent years due to its high bandwidth and data rate capabilities. FSO can provide promising wireless communication, which can support the rapid growth of different cloud applications such as internet and cell phones [1]–[4]. FSO could provide transmission with as high data rates as in optical fibers. However, in FSO, the data is transmitted via light over a FSO channel without cables (as in optical fibers). Consequently, FSO can create more flexible networks than optical fibers leading to a significant decrease in power consumption. Moreover, it is easier and cheaper to install new FSO networks than optical fiber networks. In comparison to radio frequency (RF), FSO is better in many ways and faster than RF systems. In FSO there is no need for a spectrum license as in RF systems, and the data is transferred over line of sight (LOS) so there is no need to use complicated security systems as in RF. Consequently, FSO is more secure than RF and is resistant to RF interference [5], [6]. However,

in FSO, the data is often transmitted through a turbulent FSO channel. In the turbulent channel there are random changes in the refractive index resulting in random refractions [3], [4]. FSO transmitted signals are very sensitive to these fluctuations that lead to changes in the amplitude and the phase of the received signal. This can affect the performance of FSO systems and lead to significant increases in the values of the bit error rate (BER), which can limit the implementation of FSO communication systems in real environments such as in data centers. This is because recovering the transmitted data at the receiver depends on prior knowledge of the encoder and decoder, and accurate knowledge of the channel state information (CSI). Using deep learning (DL) algorithms for recovering the transmitted data in FSO communication could be an efficient solution to employing FSO in turbulent channels, and thus permit the use of FSO without the need for any prior knowledge of the turbulent channel. The deep neural network (DNN) is one of the most commonly used algorithms in DL. DNN can optimize the performance of the entire system and learn the relationship between the input and the output of a system through training and learning processes.

Recently, researchers have widely demonstrated DL in many areas, such as in computer vision and speech recognition [7], [8]. They have also succeeded in applying DL in different areas of wireless communication systems, for encoding, decoding, modulation recognition, and channel estimation [9]–[11]. In [12], researchers proposed the use of a DL autoencoder to replace both the transmitter and the receiver of the communication system. Also, in [13], the use of DL to reduce the peak to average power ratio problem in orthogonal frequency division multiplexing (OFDM) was proposed. In [14], the authors succeed in implementing DL signal detection and channel estimation for OFDM systems. The use of DL in wireless communication systems could increase their performance as DL has algorithms and tools that enable learning different complicated models. These try to optimize the performance of the entire communication system by training and learning processes without the need to use any prior knowledge of mathematical models or parameters of the channel [15]. In addition, DL has been implemented for different applications in optical communication systems [16], such as reducing the computational complexity in different optical communication tasks [17], atmospheric turbulence detection and adaptive techniques for orbital angular momentum based FSO communication [18], mitigating fiber induced nonlinearity [19], modulation format identification in digital coherent receivers [20], and optical performance monitoring [21]. In [22], DL was used as a detection technique in FSO communication. In [18], DL was demonstrated for the detection and adaptive demodulation of orbital angular momentum based FSO communication. In [23], sensor less FSO communication was corrected using DL. In [24], the researchers used DL as a solution for an imperfect channel state information problem in correlated FSO communication channels. In the case of weak atmospheric turbulence with perfect channel state information, DL has been applied to achieve the same performance as the maximum likelihood detector. In channels without correlation, DL enabled better performance compared to that of a maximum likelihood detector. In [25], channel estimation in FSO communication was carried out using DL. However, only a few works recently suggested using DL in FSO communication.

In this paper, we suggest a new detection method for on-off-keying (OOK) modulated signals in FSO communication systems using DL models which can effectively replace the maximum likelihood (ML) decoders. We propose doing so over different FSO turbulent channels. We built two different decoders using DL. In the first decoder, we used fully connected (FC) layers. In the second decoder, we used fully convolutional neural networks (FCNN) with concatenation of memory from previous layers. In order to check our modules, we generated random data bits and OOK modulation. We then transmitted the modulated data bits via light through atmospheric turbulent channels. We used channels with weak, moderate, and strong turbulence and we compare between the performance of our models and the performance of the ML detector with perfect CSI and with the performance of the traditional OOK decoder with a fixed threshold. The results indicate that our DL decoders performed approximately like the ML decoder with perfect CSI when the channel is described by weak atmospheric turbulence. In the case of moderate and strong turbulence our decoders offer even slightly better performance, but the advantage of our decoder in this case is that it can predict the data faster than the ML decoder with perfect CSI, and with less computational complexity. When the channels are characterized by strong or moderate turbulence, we obtained improvements in the performance of the detection of OOK modulated signals compared to the OOK decoders with a fixed threshold. Moreover, our decoders led to a significant decrease in the power consumption of the detection method of OOK, and our models can effectively replace the ML decoders with perfect CSI. In addition, we succeeded in recovering the transmitted data irrespective of the strength of the turbulence. Even in the case of strong turbulence we were able to improve the performance and decrease the BER. We obtained low BER with lower signal to noise ratio vales compared to the OOK decoders with a fixed threshold. This leads to a decrease in the power consumption of the system, because we can decrease the signal to noise ratio of the transmitted signal and transmit with less energy, while obtaining better BER. Consequently, in our models a large amount of data was not generated, which is a critical problem in different DL systems consuming time and high computational power. It was sufficient to generate a small number of training data with a size of 5000 vectors of input/output data. The training time in the decoders was therefore less than in the existing DL models. After the training process, the weights of our DL system are saved, as a result of which the online transmitted data can be predicted according to the saved weights, which can recover faster than the prediction time of a regular decoder. Accordingly, the performance of detected OOK modulated signals through different turbulence channels is improved.

The novelty of our work lies in the DL models enabling the use of FSO wireless communication in data center environments with turbulent channels where the turbulence is unknown. Turbulence may arise in data centers themselves because of heating, in addition to natural turbulence in the atmosphere. We show that DL makes it possible to communicate reliably through turbulence, including heavy turbulence, using normal transmitter power without the need for previous knowledge about the communication channel. Also, our decoders are efficient and thus can also be useful in channels without turbulence, because our models succeed to exhibit good performance and decreased power consumption. Our DL decoders could replace the OOK decoders that use a fixed threshold, which consume a lot of energy when the channels are with moderate or strong turbulence. In addition, our DL decoders can replace the state of the art ML with perfect CSI decoders, because ML requires accurate CSI and for perfect channel estimation the probability density function (PDF) needs to be stationary over the window periods. However, this requirement does not hold in non-stationary channels such

as turbulent media. Thus, ML becomes essentially non-ideal and non-optimal. On the other hand DL does not make any prior assumption, and is data-driven so it is fully adaptive. DL may adapt faster and thus may exceed ML in some cases such as in strong turbulence. Consequently, our DL decoders could replace the state of the art ML decoders and solve the above problems, while at the same time succeeding to obtain performance at least as good as ML decoders, especially through turbulence.

The rest of the paper is organized as follows. In section II, the FSO turbulent channel and the potential of using DL in FSO atmospheric turbulent channels are described. In section III, the DL detection models for OOK in FSO turbulent channels are presented. Simulation results are provided in section IV. In section V, we conclude and summarize the study.

II. FSO TURBULENCE CHANNEL

FSO communication is an optical communication technology that transmits data via light through free space using intensity modulation (IM). The transmitted data propagates through a turbulent channel with additive white Gaussian noise (AWGN) [3], [4]. At the receiver, the data is received via a photodetector (PD) and is detected using direct detection (DD), as in Fig. 1:
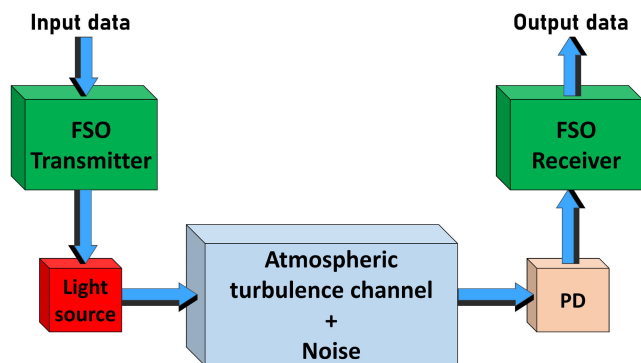

FIGURE 1. Block diagram of IM/DD FSO communication system.

We assume that the channel is memoryless and stationary, and exhibits slow fading. The received signal can be described by the basic channel model [26], [27]:

$$y_k = \eta h x_k + noise \tag{1}$$

where η is the responsivity of the PD (measured in V/W), h is the channel state, which includes attenuation due to atmospheric turbulence, and is equal to the channel intensity at that time. It is affected by distortions due to atmospheric turbulence generated by random changes in the temperature and pressure of the atmosphere. Generally, turbulence is modelled by a lognormal distribution or by Gamma-Gamma random variables in the cases of weak or strong turbulence, respectively. The intensity of the transmitted bit that is modulated using OOK modulation is $x_k \in \{0, 1\}$, and noise is signal independent AWGN with zero mean and variance $N_0/2$.

In free space where the data propagates, there are random fluctuations of refractive index, and scattering by fog, clouds, etc. Equation 1 characterizes the received data affected by distortions caused by turbulence. The fluctuations in the turbulent channel can be described by the parameter C_n^2. C_n^2 is the refractive index structure coefficient that describes the fluctuations and changes in air temperature through the channel, $k = 2\pi/\lambda$ is the wave number, λ is the wavelength, and L is the distance between the transmitter and the receiver. When the transmitted FSO data propagates through the air, the fluctuations can lead to signal fading and degradation in the performance of the received signal. It is customary to divide effects of turbulence on the received FSO signal [28] into two types of fluctuations, weak and strong. Rytov variance, σ_R^2, is the parameter that determines the type of turbulence and it can be calculated according to:

$$\sigma_R^2 = 1.23 C_n^2 k^{\frac{7}{6}} L^{\frac{11}{6}} \tag{2}$$

When $\sigma_R^2 \ll 1$, the turbulence is weak. Otherwise the turbulence is strong. In the case of weak turbulence, the distribution of the intensity of the received signal passing through the weak-turbulence channel is lognormally distributed with a PDF:

$$f_I(I) = \frac{1}{2I} \frac{1}{\sqrt{2\pi \frac{\sigma_R^2}{4}}} \exp\left[-\frac{(\ln(I) - \ln(I_0))^2}{2\sigma_R^2}\right] \tag{3}$$

where I is the received signal intensity, σ_R^2 is the variance of the log amplitude of the received signal, and $\ln(I_0)$ is the average log intensity of the received signal. For longer distances, or when the fluctuations are higher and the turbulence is strong, the distribution of the received signal is Gamma-Gamma with a PDF [29]:

$$f_I(I) = \frac{2(\alpha\beta)^{\frac{(\alpha+\beta)}{2}}}{\Gamma(\alpha)\Gamma(\beta)} \cdot I^{\frac{(\alpha+\beta)}{2}-1} \cdot k_{\alpha-\beta}(2\sqrt{\alpha\beta I}) \tag{4}$$

where $\Gamma(.)$ is the gamma function, k_i is the order of the Bessel function of the second kind, and α and β can be calculated according to:

$$\alpha = \frac{1}{\exp\left[\frac{0.49\sigma_R^2}{\left(1+1.11\sigma_R^{12/5}\right)^{7/6}}\right] - 1} \tag{5}$$

$$\beta = \frac{1}{\exp\left[\frac{0.51\sigma_R^2}{\left(1+0.69\sigma_R^{12/5}\right)^{5/6}}\right] - 1} \tag{6}$$

OOK is a modulation technique that is widely used for IM/DD in FSO communication systems due to its simplicity. A schematic of a simple OOK transmitter/receiver system using amplitude shift keying (ASK) modulation is displayed in Fig. 2:

In this modulation technique, a bit "one" is modulated by the carrier frequency and represented by an optical pulse. When the bit is zero the transmitter is in mode "off" and,
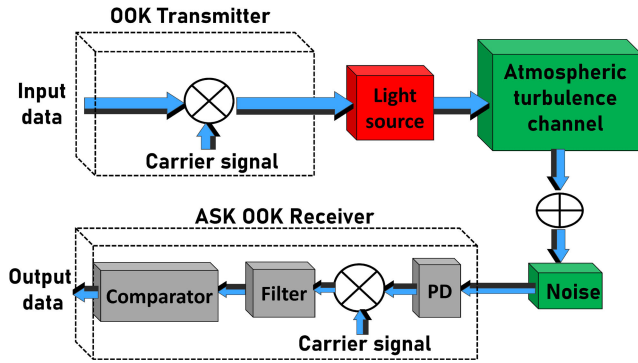
FIGURE 2. An IM/DD OOK transmitter/receiver system.

in this time interval, the transmitter is not active and does not transmit any optical power. The transmitted signal is then passed through atmospheric turbulent channel with AWGN and detected at the receiver. At the receiver, a PD detects the received power of the signal, and the signal from the PD enters a demodulator. The demodulator multiplies it by the same carrier frequency that was used in the transmitter and filters it. At the exit from the receiver, a comparator converts the analog signal to a digital signal according to a threshold that decides if the detected bit is zero or one, depending upon whether the value is less than the threshold or not. This modulation is very susceptible to noise interference because the noise affects the amplitude of the transmitted signal.

The BER calculation of IM/DD OOK modulation in FSO communication that propagates through a turbulent channel with AWGN is given by [20]:

$$P_{error} = P\left(on\right).P\left(error \mid on, I\right) + P\left(off\right).P\left(error \mid off, I\right)$$
(7)

where I is the intensity of the transmitted signal, $P\left(on\right)/P\left(off\right)$ are the probabilities of transmitting bits one or zero, and $P\left(error \mid on, I\right)/P(error \mid off, I)$ are the conditional error probabilities when the transmitting bit is one or zero. We can assume that $P\left(on\right) = P\left(off\right) = 0.5$ and the noise distribution is independent of the bit that is transmitted. The conditional bit error probability of I can be calculated according to:

$$P\left(error \mid on, I\right) = P\left(error \mid off, I\right) = Q\left(SNR.I\right)$$
(8)

where SNR is the signal to noise ratio. The average BER over the noisy channel can be calculated via the following equation:

$$BER = \int_{0}^{\infty} f_I\left(I\right).Q(SNR.I)$$
(9)

where $f_I\left(I\right)$ is the PDF of the received signal at the receiver. We mentioned above that weak turbulence is lognormally distributed and strong turbulence is Gamma-Gamma distributed. In both cases, $f_I\left(I\right)$ depends on the scintillation index parameter and inversely affects the BER. When the scintillation index parameter is increased according to equations 3 and 4, the

BER increases. In order to achieve lower bounds of BER and good performance in the case of higher values of the scintillation index parameter, the SNR is increased and more power is transmitted. Alternatively, other mitigation techniques can be applied. In some cases, it is difficult to realize FSO communication and to achieve these lower bounds of BER, or the system may consume too much energy. In OOK modulation, in order to achieve good performance one may employ ML with perfect CSI. These decoders are complicated and the receiver needs to have accurate knowledge of the instantaneous CSI. The receivers use thresholds in the detection of the recovered data in order to achieve optimal performance. The receiver also needs to know the accurate CSI to adjust the threshold, while in practical systems this parameter is unavailable. Hence, there is a significant demand to find an efficient solution to these problems in order to benefit from the advantages of FSO, and to enable the use of such communication in turbulent channels. In the next section, DL is proposed and demonstrated to be able to overcome the above problems.

III. PROPOSED DL DETECTION MODELS FOR FSO
A. AN INTRODUCTION TO DL

DL is a neuron model type learning system like a black box with an input and output. DL is used to improve the performance of a system from the experience gained via a training process, until minimum loss between the output of the system and the original data is obtained. The input to this system is corrupted data and the output data is the original data before the corruption. At the entrance to the system, the input data is mapped to a number of nodes called the input layer. Values of these nodes propagate through a graph which contains a number of layers, each of which contains nodes. Values of the nodes in each layer are determined by a specific mapping function called an activation function. At the beginning of the training process, DL sets random values to the weights and the biases, and then tries to change these values according to derivations in a number of iterations until minimal loss between the output of the system and the original data is obtained. In other words, we try to recover the corrupted input data with minimum loss. DL succeeds in maximizing the performance of a system without prior programming. Because of these advantages, in recent years researchers have widely used DL in many fields, including computer vision, speech recognition, and more. In these fields, DL succeeded in improving system performance. There exists some similarity between wireless communication and fields like speech recognition: in the two systems data bits are generated, transmitted through a channel and arrive at the receiver, which attempts to detect the original data with minimum loss. Hence, researchers started to apply DL in different fields of wireless communication systems, as mentioned above. The authors of [12]–[14] suggested using DL for signal detection in order to replace the receiver in OFDM wireless communication systems. We believe that

if researchers succeeded in using DL in signal detection and successfully replaced the receiver in wireless communication systems, then it is also possible to use DL for signal detection in FSO communication systems.

B. PROPOSED DL DETECTION MODELS

In this section, we present a new DL detection method for OOK with IM/DD modulation to enable the use of FSO communication in different turbulent channels without the need of prior channel. The aim of our DL detection models is to replace the optimal ML with perfect CSI detection OOK model and to replace the OOK decoders that use the fixed threshold, in order to receive modulated OOK data, and to recover the original data with minimum loss. The proposed DL models work in two phases. The first phase is called the training process. In this process, a data set enters the DL system, and it trains and learns in a number of iterations to recover the original data. When the system finishes this process and manages to obtain minimum loss, it saves the weights of the system. After the training process the system starts an online process, where the DL system can receive online OOK transmitted data with noise after passing through the turbulent channel. The DL model makes predictions of this data according to the weights saved at the end of the training process. The online process is expected to run faster because the weights have already been determined and saved.

In our work, we suggest two different DL detection models. In the first model, we use FC networks. In the second model we use FCNNs with concatenation of memory from previous layers. Schemes of the different DL models that we built are presented in Figs. 3 and 4.
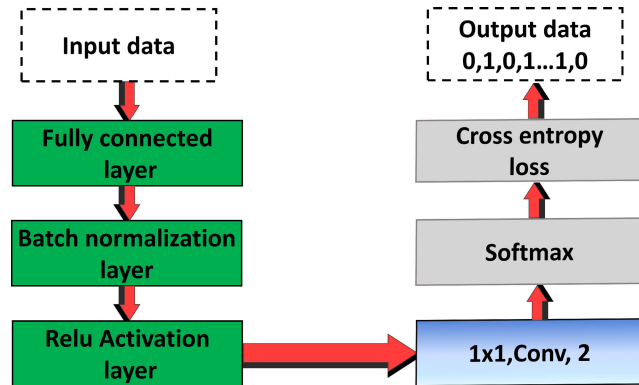

FIGURE 3. OOK decoder 1 using FC network.

In our DL models, we used a Relu activation function: $f_{Relu}(x_i) = \max(0, x_i)$ after each internal layer and the last layer is a convolutional layer with two filters of size 1×1. This layer is followed by a softmax activation layer:

$$f_{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{k} e^{x_j}} \qquad (10)$$

that converts the values of the output data from this layer to probabilities from values 0 to 1. At the end, we used the cross-entropy loss function given by equation 11 to measure the
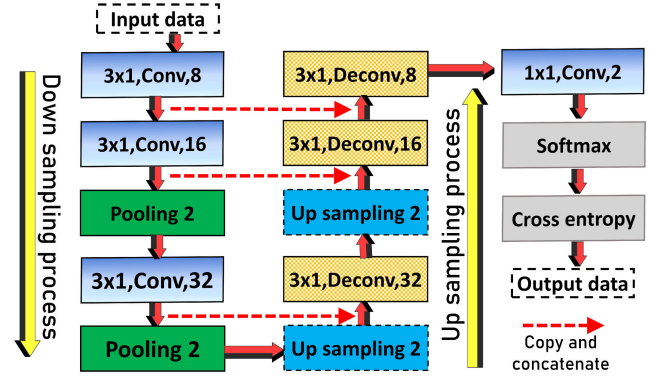

FIGURE 4. OOK decoder 2 using FCNN + concatenation of memory.

difference between two probabilities: the probability of the original data bits p and the estimated output probability q of our DL system. The distance between the output of the DL system after the softmax at the last layer, and the original data bits needs to be minimum by cross-entropy loss.

$$H(p, q) = -\sum_{x_i} p(x_i)\log(q(x_i)) \qquad (11)$$

In the first model, we used one FC layer, which means that all the input data are connected to all the nodes in this layer (see Fig. 3). After the FC layer, a convolutional layer was used with 2 filters, each one with a size 1×1. At the output of this layer we used a softmax activation function, followed by a binary cross-entropy loss function. In the second model, FCNN with concatenation of memory from previous layers was used. The concept of this model is taken from [31]–[33]. In [31], the authors suggested using FCNN for image segmentation, which means the detection, for each pixel in the image, if it is background or foreground. In [33], the authors extended this work and improved system performance by suggesting adding memory from previous layers to detect more sophisticated features.

FCNN includes two processes, namely down sampling and up sampling. The down sampling process comprises a number of convolution and pooling layers, and the up sampling process performs the inverse processes, comprising a number of up sampling and deconvolutional layers. The down sampling process is used to detect high resolution information and features in the image, in other words to extract the data. This recovers the lost information due to the convolutional and pooling layers and obtains the precise information and localization of the extracted data by an up sampling process. In the proposed model, we apply the same concept as in [31]–[33], but we modify for our problem. The purpose of our models is to insert corrupted OOK modulated data and perform semantic segmentation for each output bit to determine if it is 0 or 1. The scheme of our FCNN model is presented in Fig. 4. The input data to our DL system passes two processes of down sampling and up sampling. In the down sampling process, we used a convolutional layer with 8 filters, each one with a size 3×1. Then we used another 2 convolutional layers which duplicated the number of the filters to 16 and 32.

After each convolution layer, we used pooling layers. In the up sampling process, we used the inverse process that we used before in the down sampling process. After the up-sampling process, we used convolutional layers of two classes followed by a softmax activation function. Then, the cross-entropy loss function is used according to equation (11) to calculate the minimum loss between the input and the detected data bits. The two proposed models received the same modulated OOK data after passing through turbulence channels with AWGN, and the output of these models is a vector of data bits recovered through the cross-entropy loss function. In the next section, we present the simulation results that we perform by our different DL models.

IV. SIMULATION RESULTS

In order to check the performance of our suggested DL models in this article, we need to generate sets of input and output data, train and test our models, and then compare between the performance of our models and the performance of the ML with perfect CSI detection model and the OOK with fixed threshold detection method. For this, using MATLAB software we generate two datasets of 5,000 and 10,000 vectors of random data bits, each one with a size of 512 bits. We modulated each vector of data by OOK modulation and transmitted it across different strength FSO turbulence channels with AWGN. The receiver received the transmitted data with the noise and recovered the original data bits by two detection methods, OOK with fixed threshold and the state of the art ML with the perfect channel estimation OOK detection method. The input of our DL models is the received modulated data with noise that arrived at the receiver, and the output is the original data bits that we generated. We built our DL decoders using Tensorflow software and we ran our simulations on a computer with CPU: Intel core i7-7500 2,7 GHz. The proposed system of our DL models is shown in Fig. 5.
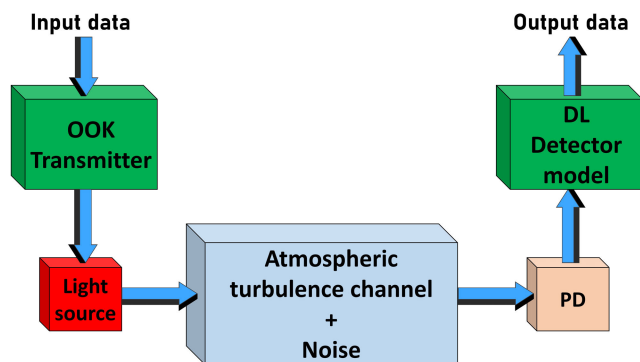

FIGURE 5. Proposed system model.

We performed simulations for different strengths of turbulent channels; weak, moderate, and strong. The strength values of the turbulent channels that we used are presented in Table 1, and the hyperparameters of the DL models that we used are presented in Table 2. These were chosen based on our previous experience with DL.
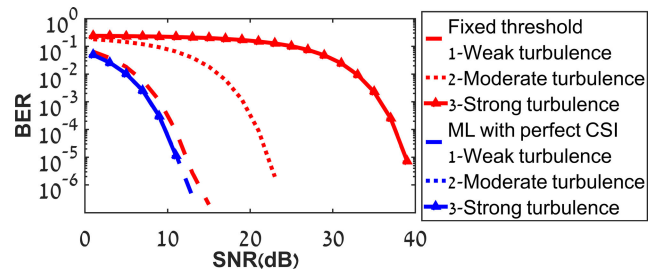

FIGURE 6. BER performance against normalized SNR for OOK with fixed threshold detection model and with ML with perfect channel estimation detection method across all turbulence channels.

TABLE 1. Strength parameters for the different turbulent channels.

	Turbulence regime	σ_R^2
Channel 1	Weak	0.1
Channel 2	Moderate	1.6, α=4, β=1.9
Channel 3	Strong	3.5, α=4.2, β=1.4

TABLE 2. Hyperparameters of the DL.

Hyperparameter	Value
Number of iterations	50
Activation function	Relu
Batch size	10
Loss	Softmax cross-entropy
Optimizer	Adam
Learning rate	0.001

The BER performance of the data that we generated by MATLAB are calculated by equation (12) and are presented in Fig. 6.

$$BER = \frac{\text{Number of error bits}}{\text{Number of total transmitted bits}} \quad (12)$$

The detection of the received data bits in the first detection method is calculated according to the threshold. If the value is higher than the threshold then the detected bit is 1, and otherwise it is 0. To adjust the optimum threshold, previous knowledge of the CSI is required. However, recovering the received data bits in the ML with perfect channel estimation requires pilot transmission of data, which reduces the data rate of the FSO transmission system.

In Fig. 6 we present performance against normalized SNR for OOK with a fixed threshold (red curves), and with ML with perfect CSI detection which (blue curves). However, the blue curve for ML represents 3 situations which yield essentially the same results; therefore, only one curve is shown. We can see in Fig.6 that as the turbulence strength increases, the BER for the detection method with a fixed threshold of OOK also increases. Thus, for channels with strong turbulence it is necessary to transmit the modulated signal with more power than for a weak turbulence channel or to use other mitigation techniques. For example, to achieve a BER lower than 10^{-5} for channel 1 with weak turbulence, it is necessary to transmit with a SNR = 13 dB. In order

to achieve the same BER in channel 3 with strong turbulence, it is necessary to transmit with SNR = 39 dB, which is 26 dB greater than that required for the weak turbulence channel. In addition, however, the performance of ML with a perfect CSI decoder, is better than the performance of the first decoder with the fixed threshold, but to achieve this performance, the receiver needs to have previous knowledge of the CSI obtained from the transmission of pilot data, which decreases the efficiency of the bandwidth and leads to a decreased data rate of the system. However, in some cases, when the channel is non-stationary or is strongly turbulent, it is difficult to know the accurate CSI in order to implement FSO, which limits the implementation potential of this technique.

The above problems can be substantially minimized by replacing the ML with perfect CSI detection and the decoders with a fixed threshold using the proposed DL models. In order to check the BER performance of the proposed DL models, two datasets of the input/output data are trained. The training of the data was carried out using Tensorflow software. During the training process, the DL system tries to learn the weights and recovers the detected data bits with minimum loss. Fifty iterations were sufficient for obtaining minimum loss. Then, after the training process, the system can receive online transmitted data and detect it. Comparison between the BER performance of our FC DL model, the ML with perfect CSI detection method, and the detection method with the fixed threshold across the different turbulent channels with the first data set are presented in Figs. 7a-c.

In Figs. 7(a)-(c), we compare between the results of the conventional detector with a fixed threshold, ML with perfect CSI detector, DL FC model 1, and FCNN with memory model 2 for the three different turbulent channels. The blue curve presents BER performance of FC model 1, the green curve presents the results of the conventional detector with a fixed threshold, the red curve shows the performance of the ML with perfect CSI detector, and the black curve shows the results for the DL FCNN with memory model 2. Across all the different turbulent channels, the proposed DL models present better performance and energy consumption than those of the conventional detection method with a fixed threshold. The performance results are close to those of the ML detector with perfect CSI. For example, when the channel is turbulent the results of the ML detector and our models are very close, but when the channel is with moderate or strong turbulence, our models display an improvement over ML performance by a few dB. The results for FC (model 1) and FCNN with memory (model 2) are very similar and consume less energy than the regular detection method with the fixed threshold and are also very close to the results of the ML detector model. For example, when the channel is with strong turbulence, to obtain BER = 10^{-5} in the conventional detector with a fixed threshold, it is necessary to transmit with SNR = 39 dB. However, in the FC model 1 and FCNN with memory model 2 cases, it is sufficient to transmit with SNR = 8 dB and SNR = 9 dB, respectively. These levels are
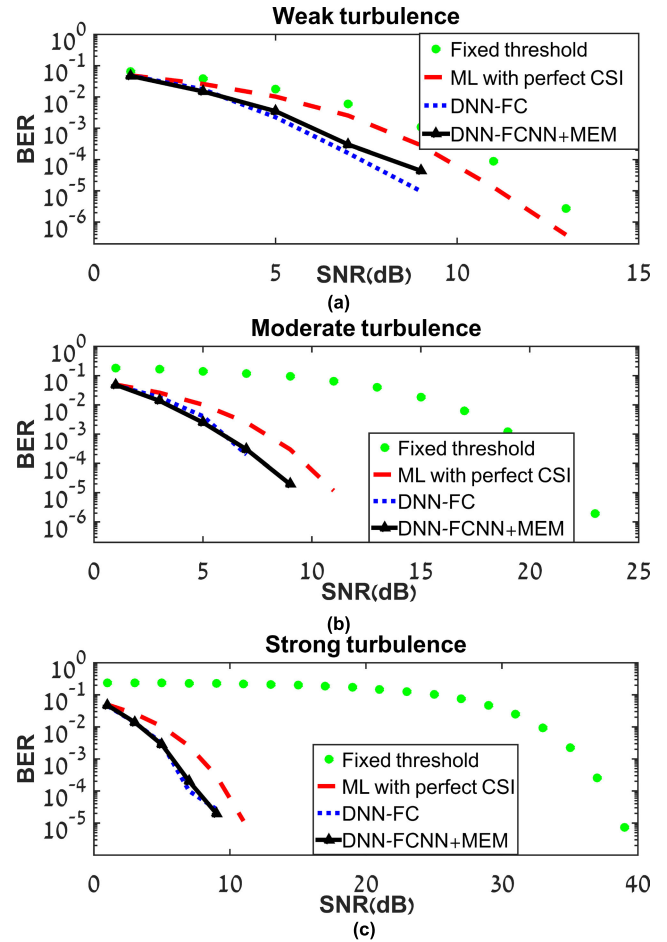

FIGURE 7. BER performance against the normalized SNR for conventional detector with fixed threshold, ML with perfect CSI detector, DNN FC model 1, and DNN-FCNN with memory model 2.

approximately 30-31 dB less than the required values for the regular detector with a fixed threshold, and 1-3 dB less than the required SNR in the ML with perfect CSI detection case. In our case we get performance very close to the ML performance, and even slightly improved (by 1-3 dB) since DL is more robust to variations in the AWGN channel models than ML. In our case, ML knows the channel coefficient of the turbulent channel, while DL considers the whole system as a black box channel coefficient of the turbulent channel with the AWGN. Since DL tries to minimize total system loss, we get slightly better performance. Another important thing in DL is how to set the values of the hyperparameters of the network. In order to get better performance it is very important correctly set the hyperparameters. When the hyperparameters are tuned properly, the network can learn more complex relationships. Any small changes in these parameters affects the outcome and leads to worse performance. Further, in DL we could go deeper and implement many layers, so that DL can yield more complicated features.

In Figs. 8(a)-(c), the results of the proposed FC model are presented with two data sets of sizes 5,000 and 10,000. The performance of the first data set with size 5,000 is shown in black, and results of the second data set are presented in red.
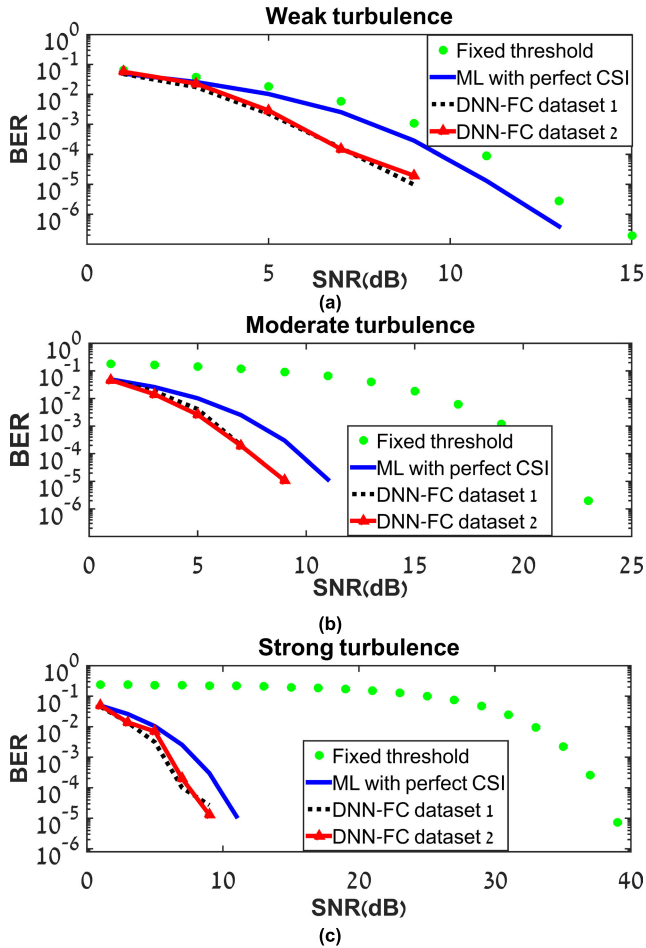
FIGURE 8. BER performance against normalized SNR for FC model datasets equal to 5,000 and 10,000.

FIGURE 9. Number of FLOPs for the different models.
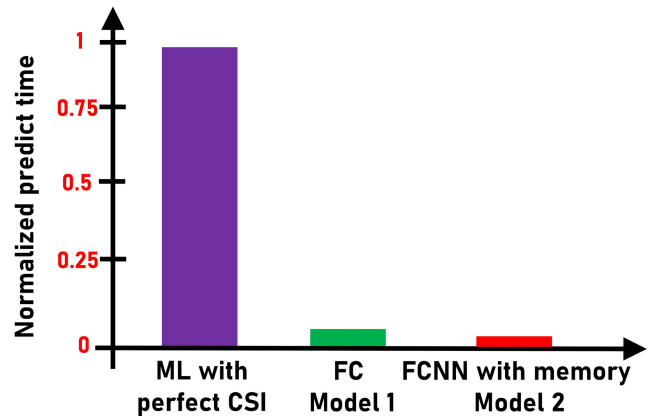
FIGURE 10. Normalized detection time consumption for the different models.

In Figs.8a-c we see that the results of the two data sets with different sizes are very close and the small data set of size 5,000 is sufficient to obtain good results and decrease by half the training time compared with the long dataset. The two data sets yield good performance, better than that of the regular detector with the fixed threshold and slightly better than the performance of the ML decoder with perfect CSI and are very similar. For example, in the results of the two data sets we can obtain a BER of less than 10^{-5} with a SNR of approximately 4 dB, 13 dB, and 30 dB less than in the three different turbulent channels, respectively, compared with the fixed threshold detection method, and approximately 1-3 dB less than with the ML with perfect CSI detection method.

We calculate the complexity of the proposed models in terms of amount of floating point multiplication adds (FLOPs) and detection time consumption (see Figs. 9 and 10).

In Figs. 8a-c, it is shown that the performance results of the FC (model 1) and FCNN (model 2) cases are very close, but according to Fig. 9, the number of FLOPs in FCNN with memory model 2 is seen to be less than for the FC model. This is because in FCNN the nodes in each layer are not connected to all the nodes in the next layer, which leads to
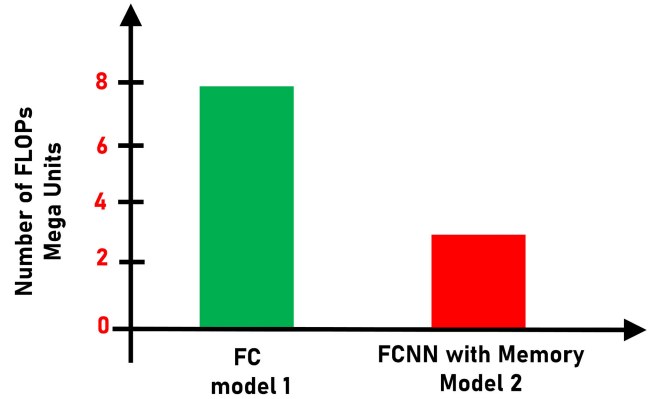
reduced calculations. Moreover, the detection time after the training process in the models for one input data set was less than 0.01 times the detection time in the ML with perfect CSI detection method. Therefore, the proposed models succeed to recover the detected data with improved performance and speed in the OOK detection model than ML with perfect channel estimation detectors. Our models are simple and very easy to use and, at the same time, succeed in achieving similar performance to the ML with perfect CSI detection methods.

V. SUMMARY AND CONCLUSION

The novelties of the methods presented here are the effectiveness of using DL for signal detection of OOK modulated data over different FSO turbulent channels in terms of: improving the performance, decreasing the prediction time, and reducing the energy consumption. In such situations, ML is non-ideal and non-optimal because the channel is not stationary. We built two different models of DL; in the first model we used FC neural networks and in the second model we used FCNN with concatenation of memory from previous layers. We tested our models in two phases. In the first phase we trained our models offline using OOK modulated data that was received after passing through different turbulence strength channels with noise. During this process, the models learn the weights of the system. In the second phase, the

system received online transmitted modulated OOK data with noise and recovered the original data bits.

We compared between the performance of our suggested DL models and the ML with perfect channel estimation OOK detection method and with the fixed threshold detection method. In the simulation results, we show that the use of DL for signal detection of OOK has many advantages when the FSO channel has strong turbulence. DL successfully recovered the original data bits with a significant improvement in BER performance compared with the fixed threshold detection method and performed slightly better than the state of the art ML with perfect CSI decoder method. DL was able to detect the data and learn the channel despite the turbulence, no matter if the turbulence was strong or weak. For example, over a strong turbulence channel with $\sigma_R^2 = 3.5$, in order to obtain a BER of less than 10^{-5}, the DL models succeeded in decreasing the required SNR by 30 dB compared with the fixed threshold detection method, and by 1-2 dB compared with ML with perfect CSI. The DL models were able to detect the data with similar BER performance for the different turbulence levels, with lower BER than in the fixed threshold detection method. We succeeded in obtaining the same BER performance with approximately the same SNR = 8 to 10 dB, which is less than the required SNR in the fixed threshold detection method, which was 14 dB, 25 dB, and 39 dB for the three different turbulent channels that we used with $\sigma_R^2 = 0.1$, 1.6 and 3.5. In addition, our DL decoders performed slightly better than the state of the art ML with perfect CSI decoder. If we use DL for signal detection of OOK in FSO, it is possible to exploit the many advantages of this technology, such as communicating during turbulence. Also, in cases of unknown channel parameters, such as in a fast changing channel, we can use FSO and transmit with less energy, while obtaining BER levels characteristic of an environment with weak turbulence or a deterministic channel. Moreover, after the training process, our models work faster than the ML detection method. Hence, the advantage of using FCNN is that we can insert any non-specific size of input data, although the input data should be equal to or a multiple of 512.

ACKNOWLEDGMENT

REFERENCES

[1] Lund, Denise, "Worldwide and regional Internet of Things (iot) 2014-2020 forecast: A virtuous circle of proven value and demand," in *Proc. Int. Data Corp. (IDC)*, vol. 1, 2014, pp. 1–4.

[2] I. Yaqoob, E. Ahmed, I. A. T. Hashem, A. I. A. Ahmed, A. Gani, M. Imran, and M. Guizani, "Internet of Things architecture: Recent advances, taxonomy, requirements, and open challenges," *IEEE Wireless Commun.*, vol. 24, no. 3, pp. 10–16, Jun. 2017.

[3] H. Kaushal, V. K. Jain, *Free Space Optical Communication*, vol. 18. New Delhi, India: Springer, 2017.

[4] N. Blaunstein, S. Arnon, and A. K. N. Zilberman, *Applied Aspects of Optical Communication and LIDAR*. Boca Raton, FL, USA: CRC Press, 2009.

[5] A. Malik and P. Singh, "Free space optics: Current applications and future challenges," *Int. J. Opt.*, vol. 2015, pp. 1–7, Oct. 2015.

[6] S. Chaudhary and A. Amphawan, "The role and challenges of free-space optical systems," *J. Opt. Commun.*, vol. 35, no. 4, pp. 327–334, Jan. 2014.

[7] E. Charniak, *Introduction to Deep Learning*. Cambridge, MA, USA: MIT Press, 2019.

[8] A. Fehske, J. Gaeddert, and J. H. Reed, "A new approach to signal classification using spectral correlation and neural networks," in *Proc. 1st IEEE Int. Symp. New Frontiers Dyn. Spectr. Access Netw.*, Nov. 2005, pp. 144–150.

[9] R. Azzouz, "Modulation recognition using artificial neural networks," in *Automatic Modulation Recognition of Communication Signals*. Boston, MA, USA: Springer, 1996, pp. 132–176.

[10] M. Ibukahla, J. Sombria, F. Castanie, and N. J. Bershad, "Neural networks for modeling nonlinear memoryless communication channels," *IEEE Trans. Commun.*, vol. 45, no. 7, pp. 768–771, Jul. 1997.

[11] X. Li, F. Dong, S. Zhang, and W. Guo, "A survey on deep learning techniques in wireless signal recognition," *Wireless Commun. Mobile Comput.*, vol. 2019, pp. 1–12, Feb. 2019.

[12] T. O'Shea and J. Hoydis, "An introduction to deep learning for the physical layer," *IEEE Trans. Cognit. Commun. Netw.*, vol. 3, no. 4, pp. 563–575, Dec. 2017.

[13] M. Kim, W. Lee, and D.-H. Cho, "A novel PAPR reduction scheme for OFDM system based on deep learning," *IEEE Commun. Lett.*, vol. 22, no. 3, pp. 510–513, Mar. 2018.

[14] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.

[15] J. Sjöberg, Q. Zhang, L. Ljung, A. Benveniste, B. Delyon, P.-Y. Glorennec, H. Hjalmarsson, and A. Juditsky, "Nonlinear black-box modeling in system identification: A unified overview," *Automatica*, vol. 31, no. 12, pp. 1691–1724, Dec. 1995.

[16] C. M. Bishop, *Pattern Recognition and Machine Learning*. Cham, Switzerland: Springer, 2006.

[17] J. M. Zurada, *Introduction to Artificial Neural Systems*, vol. 8. St. Paul, MN, USA: West, 1992.

[18] J. Li, M. Zhang, D. Wang, S. Wu, and Y. Zhan, "Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication," *Opt. Express*, vol. 26, no. 8, pp. 10494–10508, 2018.

[19] F. N. Khan, "Joint OSNR monitoring and modulation format identification in digital coherent receivers using deep neural networks," *Opt. Express*, vol. 25, no. 15, pp. 17767–17776, Jul. 2017.

[20] J. Thrane, J. Wass, M. Piels, J. C. M. Diniz, R. Jones, and D. Zibar, "Machine learning techniques for optical performance monitoring from directly detected PDM-QAM signals," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 868–875, Feb. 15, 2017.

[21] Y. Huang, L. Gutterman, P. Samadi, P. Cho, B. Samoud, W. Ware and K. Bergman, "Dynamic mitigation of EDF Apower excursions with machine learning," *Opt. Express*, vol. 25, no. 3, pp. 2245–2258, 2017.

[22] C. Zheng, S. Yu, and W. Gu, "A SVM-based processor for free-space optical communication," in *Proc. IEEE 5th Int. Conf. Electron. Inf. Emergency Commun.*, May 2015, pp. 30–33.

[23] Z. Li and X. Zhao, "BP artificial neural network based wave front correction for sensor-less free space optics communication," *Opt. Commun.*, vol. 385, pp. 219–228, Feb. 2017.

[24] M. A. Amirabadi, "A deep learning based solution for imperfect CSI problem in correlated FSO communication channel," 2019, *arXiv:1909.11002*. [Online]. Available: http://arxiv.org/abs/1909.11002

[25] M. A. Amirabadi, M. H. Kahaei, S. A. Nezamalhosseini, and V. T. Vakili, "Deep learning for channel estimation in FSO communication system," *Opt. Commun.*, vol. 459, Mar. 2020, Art. no. 124989.

[26] H. E. Nistazakis, G. S. Tombras, A. D. Tsigopoulos, E. A. Karagianni, and M. E. Fafalios, "Average and outage capacity estimation of optical wireless communication systems over weak turbulence channels," in *Mosharaka Int. Conf. Commun., Propag. Electron., MIC-CPE*, Jan. 2009, pp. 1–4.

[27] H. E. Nistazakis, E. A. Karagianni, A. D. Tsigopoulos, M. E. Fafalios, and G. S. Tombras, "Average capacity of optical wireless communication systems over atmospheric turbulence channels," *J. Lightw. Technol.*, vol. 27, no. 8, pp. 974–979, Apr. 2009.

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

[29] L. C. Andrews and L. Y. Ronald, *Laser Beam Scintillation With Applications*, vol. 99. Bellingham, WA, USA: SPIE, 2001.

[30] S. Navidpour, M. Uysal, and M. Kavehrad, "BER performance of free-space optical transmission with spatial diversity," *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2813–2819, Aug. 2007.

[31] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.

[32] N. Jmour and S. Zayen, "Convolutional neural networks for image classification," in *Proc. Int. Conf. Adv. Syst. Electr. Technol.*, Mar. 2018, pp. 397–402.

[33] R. Olaf, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–271.

LAIALY DARWESH (Member, IEEE) received the B.S. degree in communication systems engineering and the M.S. degree in communication systems engineering from the Ben-Gurion University of the Negev, Beer-Sheva, Israel, in 2010 and 2013, respectively, where she is currently pursuing the Ph.D. degree in electrical and computer engineering.

Since 2016, she has been a Research Assistant with the Ben-Gurion University of the Negev. From 2013 to 2016, she was an Application Developer for Android, Sadeel Software, Omer, Israel. From 2010 to 2013, she was a Developer with Ness Technologies, Beer Sheva. Her research interests include deep learning for free space optical communication in data center environments. In 2018, she was awarded an Outstanding Doctoral Fellowship for the Arab society and minorities in Israel.

NATAN S. KOPEIKA (Life Senior Member, IEEE) was born in Baltimore, in 1944. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Pennsylvania, in 1966, 1968, and 1972, respectively. He and his family moved to Israel, and he began his career at the Ben-Gurion University of the Negev, in 1973. He Chaired the Department of Electrical and Computer Engineering for two terms (1989–1993) and was named as a Reuven and Francis Feinberg Professor of Electrooptics, in 1994. He was the first head of the Department of Electrooptical Engineering, from 2000 to 2005, which grants graduate degrees in electrooptical and photonics engineering. He has published more than 200 articles in international reviewed journals and well over 150 papers at various conferences. Recent research involves development of a novel inexpensive focal plane array camera for terahertz imaging and very fast inexpensive millimeter wave wireless communication. He was awarded the JJ Thomson Prize by the IEE, U.K., in 1999. He is a Fellow of SPIE, since 2000. He is perhaps best known for his work in atmospheric optics. Other areas of research include interactions of electromagnetic waves with plasmas, the optogalvanic effect, environmental effects on optoelectronic devices, imaging system theory, propagation of light through the atmosphere, imaging through the atmosphere, image processing and restoration from blur, imaging in the presence of motion and vibration, lidar, target acquisition, and image quality in general. He is the author of the textbook *A System Engineering Approach to Imaging* (SPIE Press, first printing 1998 and second printing 2000), and is a Topic Editor for Marcel Dekker for atmospheric optics in their Encyclopedia of Optical Engineering. He has coauthored two other books.

• • •