Deep Learning Based Signal Detection for OFDM VLC Systems

Deep Learning based Signal Detection for OFDM VLC Systems

Nurul Aini Amran, Mohammad Dehghani Soltani, Mehrdad Yaghoobi, Majid Safari
Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh, EH9 3FD, UK
Email: {nurul.amran, m.dehghani, m.yaghoobi-vaighan, majid.safari}@ed.ac.uk

***Abstract*—Visible light communication (VLC) has become increasingly popular and has sparked a wide interest from various research areas. In order to fully realize the potential of VLC and to provide seamless connectivity to users, the underlying channel model of the system must be carefully understood. However, in a practical environment which considers specific geometrical configurations of the network and user behavior effects, the optical channel can often become too complex to be modelled mathematically. In this work, we apply deep learning (DL) to design an effective signal detection scheme for an indoor VLC communication system. With the aid of DL, our system is able to learn directly from the transmitted and received symbols and can reliably detect the original transmitted signals during the real time implementation with only a limited instantaneous knowledge of the channel. The simulation results confirm that our model offers very close performance to the optimal maximum likelihood (ML) detection with perfect channel state information (CSI). We also consider specific indoor environments (e.g., using a hot spot model) to confirm the robustness of the learning based schemes in different indoor scenarios.**

***Index Terms*—Deep learning, long short-term memory (LSTM), visible light communication (VLC), DCO-OFDM.**

I. Introduction

The increasing demand for wireless data has motivated both academia and industry to find alternative technologies that are capable of providing efficient, reliable and high-speed wireless data services. Visible light communication (VLC) is one of the most attractive and important representation of the future indoor optical wireless communication system [1]. VLC offers numerous advantages in comparison to RF systems which include massive and unregulated bandwidth in the visible light spectrum, high potential data rates, high energy efficiency and better security [2]. In a typical indoor environment, the information of the channel condition such as the channel gain can be exploited to improve the communication metrics such as the signal-to-noise ratio (SNR) and symbol-error ratio (SER) [3]. Most of the current research literature on VLC [4]–[6] assume general geometries of the network and the room and a rather deterministic channel model for simplicity purposes. However, for a practical implementation of VLC, we should consider the specific geometrical configurations of the network and the user behavior effects in the channel model.

Some common traditional channel estimation approaches such as least squares (LS) and minimum mean square error (MMSE) have been widely utilized in many applications. A study in [7] shows that LS estimation is preferred for its simple and easy implementation, however, since it does not require any prior channel statistics, its performance may be insufficient. MMSE estimation method is capable of offering a better performance compared to LS as it utilizes the second order statistics of the channel. Even so, it may lead to a higher computational complexity due to the inverse operations of the channel matrix. An alternative approach in acquiring a realistic optical channel can be realized by implementing deep learning (DL) algorithms in VLC. DL is capable of extracting useful characteristics of the channel (e.g., channel gain, SNR, SER, etc.) which may not be easily or directly measured using conventional approaches. Owing to this advantage, signal detection with DL can be a simple process that can be performed in real time and offers a trade-off between accuracy and complexity.

Recently, machine learning schemes such as DL have gained attraction in designing different aspects of wireless communication systems, such as modulation, positioning, resource allocation and so on [7]–[17]. In [7], the authors proposed a deep neural network (DNN)-based channel estimation technique used for learning the wireless channel characteristics in orthogonal frequency-division multiplexing (OFDM) systems, which outperformed traditional methods such as LS and MMSE. Furthermore, DL has also been used for classifying modulated signals in the physical layer as in [11]. The authors proposed an automatic modulation recognition based on the implementation of convolutional neural network (CNN) and a long short term memory (LSTM) model. In [12], the authors developed a DL-based detection algorithm for molecular communication systems. The proposed model was trained based on the known transmitted signals but in the absence of knowledge of the underlying wireless channel model. The authors in [13] employed a CNN-aided online learning model to understand the mobility patterns of users relying on analyzing the continuous mobile data streams while the authors in [14] proposed an intelligent deep LSTM aided system for predicting human mobility in order to provide an accurate handover management.

It should be noted that compared to the radio wireless systems, the VLC channel highly depends on the geometry of the indoor environment and the behavior of the user including the user mobility and random orientation of the user equipment (UE). Therefore, the LS estimation may perform poor when only partial channel state information (CSI) is available while a learning scheme can specifically adapt to the environment. Hence, we propose a deep learning-based

signal detection scheme employing LSTM neural network while taking into account some realistic characteristics of the VLC channel such as the random orientation of UE and an infinity order of the non line of sight (NLOS) channel components. We demonstrate that our DL-based scheme have the ability to indirectly model the characteristics of the channel and can efficiently recover the transmitted symbols in real time. The simulation results show that the proposed scheme can achieve a very close performance to the optimal maximum likelihood (ML) detection assuming full CSI and offers a significantly better performance compared to the classical channel estimation approach namely least squares estimation in the scenario where the system have a limited amount of channel information.

II. System Model

A. VLC Channel Model

We consider a single input/single output (SISO) configuration for the VLC system. For this configuration, only one AP transmits the data and only one PD detects the signal. LED and PD are used as optical source and detector respectively. Due to the incoherent characteristics of LEDs, the most practical modulation and down-conversion technique used for VLC is intensity modulation/direct detection (IM/DD). For a typical transmission and reception in a VLC link employing IM/DD, the signal is modulated onto the instantaneous power of the carrier. At the AP, the transmitted optical signal, $x(t)$ passes through the channel with the channel impulse response (CIR) of $h(t)$ and outputs the current, $y(t)$ at the PD. This can be described as:

$$y(t) = Rh(t) \otimes x(t) + n(t), \tag{1}$$

where R is the PD responsivity, $n(t)$ is the signal-independent additive noise at the receiver and "$\otimes$" is a convolution operator. The channel impulse response can be characterized as the combined effect of the low-pass characteristic of the LED and the optical channel. Thus, in the time domain, the CIR is given by: $h(t) = h_{led}(t) \otimes h_{opt}(t)$, where $h_{led}(t)$ and $h_{opt}(t)$ are the CIR of the LED and the indoor VLC link respectively. Therefore, the frequency response of the channel including both the LED effect and the optical channel is: $H(f) = H_{led}(f)H_{opt}(f)$, where $H_{led}(f)$ is the Fourier transform of $h_{led}(t)$ and $H_{opt}(f)$ denotes the Fourier transform of $h_{opt}(t)$.

The VLC system comprises of a single LED transmitter situated at the ceiling, facing vertically downwards and a PD receiver mounted on a UE, firstly assumed to be oriented vertically upward towards the ceiling. The VLC channel relies upon the existence of LOS and NLOS components where LOS is a phenomena when the link established between the AP and UE is direct and uninterrupted. Meanwhile, NLOS link relies upon the reflection of light and is achieved by means of reflecting surfaces (e.g., walls and furniture) in the environment. It should be noted that an infinite order of reflections should be considered in order to achieve the most accurate representation of the NLOS channel. Hence, the optical channel in an indoor VLC system can be described as:

$$H_{opt}(f) = H_{LOS}(f) + H_{NLOS}(f), \tag{2}$$

where H_{LOS} and H_{NLOS} are the frequency response of the LOS and NLOS channel components, respectively. The direct-current (DC) gain of the LOS path between the AP and the UE is given by:

$$H_{LOS}(0) = v\frac{(m+1)A}{2\pi d^2}\cos^m\phi g_f g(\psi)\cos\psi, \tag{3}$$

where A, ϕ, ψ are the detector area, radiance angle of the transmitter and incidence angle of the receiver, respectively. $m = -\frac{1}{\log_2(\cos\Phi_{1/2})}$ is the Lambertian order where $\Phi_{1/2}$ is the half-intensity angle. The optical filter and concentrator gain are denoted by g_f and $g(\psi)$, respectively. v=1 for $0 \leq \psi \leq \Psi_c$ and 0, otherwise is the visibility factor where Ψ_c is the field of view (FOV). The radiance angle, ϕ and the incidence angle, ψ can be calculated by $\cos\phi = \frac{d.\mathbf{n}_t}{||d||}$ and $\cos\psi = \frac{-d.\mathbf{n}_u}{||d||}$, where $\mathbf{n}_t = [0,0,-1]^T$ and $\mathbf{n}_u = [0,0,1]^T$ denotes the normal vectors at the AP and UE respectively whereas d is the distance vector from the receiver to the transmitter. Moreover, the symbols $.$, $||\,.\,||$ and $(.)^T$ are the inner product, the Euclidean norm operators and the transpose operator, respectively.

For the NLOS links, a high reflection order is required to obtain an accurate value of the diffuse channel components. The method described in [18] were used to consider an infinite order of reflections by calculating the channel gain in the frequency domain instead of the time domain. In order to calculate the diffuse link, the environment is segmented into a number of small surface elements which act as reflectors. Thus, the NLOS channel gain which include an infinite order of reflections is expressed as:

$$H_{NLOS}(f) = \mathbf{r}^T(f)\mathbf{G}_\rho(\mathbf{I} - \mathbf{H}_e(f)\mathbf{G}_\rho)^{-1}\mathbf{t}(f), \tag{4}$$

where the transmitter transfer function vector is denoted as $\mathbf{t}(f) = [H_{1,\mathrm{Tx}}(f), H_{2,\mathrm{Tx}}(f), ..., H_{N,\mathrm{Tx}}(f)]^T$, $\mathbf{H}_e(f)$ is the frequency-dependent transfer matrix describing the LOS transfer function between the kth reflector and the ith reflector which acts as the transmitter and receiver surface elements respectively. $\mathbf{r}$ denotes the receiver transfer function vector expressed as $\mathbf{r}^T(f) = [H_{\mathrm{Rx},1}(f), H_{\mathrm{Rx},2}(f), ..., H_{\mathrm{Rx},N}(f)]$. $\mathbf{G}\rho = diag(\rho_1, ..., \rho_N)$ is the reflectivity matrix of all N reflectors with the reflection coefficient of ρ_i for the ith reflector and $\mathbf{I}$ is the unity matrix of size NxN.

B. Random Orientation Model

In VLC, the channel gain highly depends on the random orientation of the user device. It is important to consider this factor in order to model the channel more accurately and as realistic as possible. In a recent work [19], the authors have proposed a novel model for random orientation of mobile devices which we have adopted in modelling the indoor channel of the VLC system. Based on this model, the incidence angle can be derived as $\cos\psi = a\sin\theta + b\cos\theta$ with $a = -\left(\frac{x_a - x_u}{d}\right)\cos\Omega - \left(\frac{y_a - y_u}{d}\right)\sin\Omega$ and $b = \left(\frac{z_a - z_u}{d}\right)$. Here, (x_a, y_a, z_a) and (x_u, y_u, z_u) are the location vectors of the AP and UE respectively. θ is the elevation angle between the positive direction of the Z-axis and the UE normal vector,

while Ω is the direction in which the user is facing. Then, we can directly include $\cos\psi$ into the channel gain analysis.

C. Modulation Technique

In this paper, OFDM is assumed as the modulation technique due to its resistance to inter-symbol interference (ISI). For systems employing IM/DD, the transmitted signal must be real and positive. Due to the fact that the conventional OFDM approach takes only complex and bipolar signals, a number of OFDM based modulation suitable for the optical wireless systems were introduced (e.g., DC-biased optical OFDM (DCO-OFDM) and asymmetrically clipped optical OFDM (ACO-OFDM)). Since DCO-OFDM uses all of the subcarriers to carry the data symbols, it is more bandwidth efficient compared to ACO-OFDM which only uses half of the subcarriers to carry the data symbols [20]. Hence, DCO-OFDM is considered in this work.

III. Signal Detection with Deep Learning

A. Deep Learning Method

In this paper, we propose an effective approach for detecting demodulated signals at the receiver of a DCO-OFDM system with the aid of LSTM neural network. LSTM is a recurrent neural network which is capable of learning long-term dependencies between data sequence and therefore has achieved great success in sequence prediction problems [21], [22]. LSTM networks have hidden layers which consist of memory cells controlled by 'gates' (e.g., input gate, forget gate and output gate). The gates comprise of a sigmoid function and a pointwise multiplication operation, and serve by regulating the flow of information into and out of the memory cell. Hence, they decide what new information to be input to the cell, what old information to be discarded, and what will be the updated information to be output from the cell, respectively.

The sequence of operations in LSTM at time step t can be described as [23]:

$i_t = \sigma_g(x_t W_{xi} + h_{t-1} W_{hi} + c_{t-1} W_{ci} + b_i)$,
$f_t = \sigma_g(x_t W_{xf} + h_{t-1} W_{hf} + c_{t-1} W_{cf} + b_f)$,
$o_t = \sigma_g(x_t W_{xo} + h_{t-1} W_{ho} + c_t W_{co} + b_o)$,
$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c)$,
$h_t = o_t \odot \tanh(c_t)$.

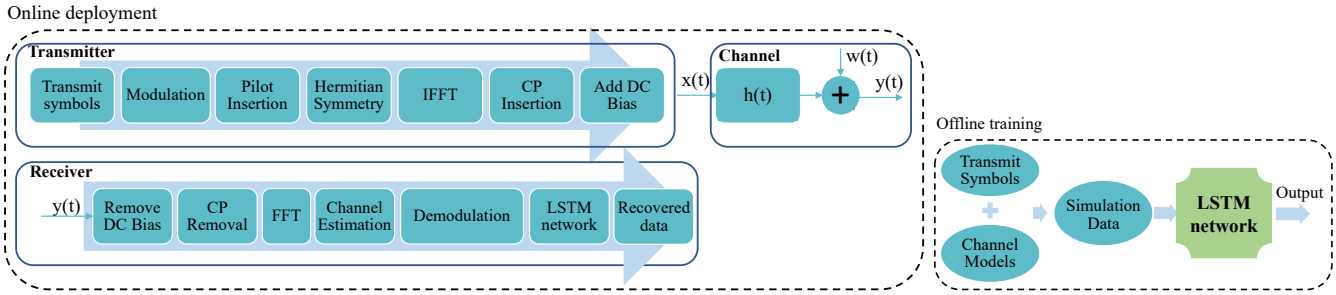
Here, x_t represents the input to the LSTM block. At the current time step t, i_t, f_t, o_t, c_t, h_t denotes the input gate, forget gate, output gate, the cell state and the output of the LSTM block, respectively. W_{xi}, W_{xf} and W_{xo} are the weights between the input layer and the input, forget and output gates, respectively while W_{hi}, W_{hf} and W_{ho} are the weights between the hidden layer and the input, forget and output gates, respectively. Furthermore, the weights between the cell state and the input, forget, and output gates, respectively are denoted by W_{ci}, W_{cf} and W_{co}. The biases of the input, output and forget gates are given by b_i, b_f and b_o, respectively. Finally, $\odot$ denotes the dot product, the gate activation function is denoted by $\sigma_g(.)$ and $\tanh(.)$ is the hyperbolic activation function. The sigmoid activation function outputs the values between 0 and 1 which decides how much information will pass through the cell gates.

B. Deep Learning Implementation in VLC

The architecture of our DL-based DCO-OFDM is illustrated in Fig. 1. As seen in the figure, the DCO-OFDM system is similar to the conventional OFDM. The main difference is that firstly, the signal being transmitted in DCO-OFDM must have Hermitian symmetry and secondly, a DC bias should be applied before transmitting the signal through the channel. This is to fulfil the requirements where the signal to be transmitted in IM/DD systems must be both real and positive. During the real-time operation, at the transmitter side, a sequence of symbols are randomly generated which then undergo modulation. Then, pilot symbols are uniformly inserted into the sequence to be transmitted. Due to the Hermitian symmetry, the transmitted signal becomes real after going through the inverse fast fourier transform (IFFT) process. This means that the input sequence to the IFFT block should be in the form of $\boldsymbol{X} = [0, X_1, \ldots, X_{\kappa/2-1}, 0, X^*_{\kappa/2-1}, \ldots, X^*_1]$, where κ is the total number of subcarriers and it should be noted that the modulated subcarriers which carry the information is only $\kappa/2 - 1$. Afterwards, in the time domain, CP is inserted and DC bias is added to make the signal positive before passing through the optical channel. After going through the channel, the received signal can be expressed as in (1). The reverse process is conducted at the receiver side where after removing the DC bias and the CP, FFT is performed to convert the signal from the time domain to the frequency domain. Hence, the received signal in the frequency domain can be described as: $Y(f) = H(f)X(f) + N(f)$ where $Y(f)$, $X(f)$, $H(f)$ and $N(f)$ are the FFT of $y(t)$, $x(t)$, $h(t)$ and $n(t)$, respectively. Finally, after demodulation, the received signal is fed as input to the trained LSTM model and the model then recovers the original transmitted data.

For the training stage, using the channel models discussed in Section II, the training data can be generated by simulation. The training data were divided into two subsets namely training set and validation set. The validation set is not used to train the network, but instead were used to monitor the validation loss during the training process in order to check if the network is overfitting the training data. In each simulation, the transmitted signal undergoes channel distortions caused by the diffuse channel components and different noise samples. This will help the DL model to have a good generalization ability during the online deployment stage. To collect the training data, a DCO-OFDM system with 128 subcarriers and QPSK modulation were considered. The transmit signal consist of 128 randomly generated symbols for 26 different SNRs ranging from 0 dB to 50 dB. The generation of the same sequence of signal is repeated but added with different sequence of independent and identically distributed (i.i.d) noise to the signal in each iteration. In addition, to consider the randomness of the device orientation, the signals were generated based on random samples of θ taken from a Laplace distribution which has a mean value of 41°.

Two scenarios were investigated; i) the location of the user is fixed, ii) the location of the user is random. For the first



(a) Online deployment (b) Offline training

Fig. 1. The structure of deep learning-based DCO-OFDM system.

scenario, the location of the user and the direction of which the user is facing are fixed (e.g., 4 users sitting around a table). For this case we consider 4 fixed locations of the user in the room with each location having a random orientation of the UE. For the second scenario, random locations, directions and UE orientations were analyzed. For this particular scenario, we consider 16 uniformly distributed locations of the user with 4 different values of Ω (e.g., 45°, 135°, 225° and 315°) for each location. It is important to note that random orientation of the user device (i.e., θ) is applied in both fixed and random cases. The transmitted and received symbols for both scenarios were collected as the training data. The weights are randomly initialised and learned through back-propagation method. During the training stage, the DL model learns to minimize the error between the output of the neural network and the original transmitted data by re-adjusting the weights using a mean square error (MSE) loss function.

The hyperparameters (e.g., the number of hidden layers, the number of neurons in each hidden layers, the learning rate, etc.) were determined by conducting several experiments on various configurations of the LSTM network. We note that the optimal performance can be achieved when our LSTM model consist of 7 layers, which are the input and output layers, three hidden LSTM layers with 50, 25 and 10 neurons respectively, one fully connected layer and one dropout layer. The dropout layer is added to help reduce the overfitting problem and the model was trained until the validation loss stops decreasing or becomes larger than the previous minimum value. Furthermore, we simulate this problem using CNN where we set the number of convolutional layers to be similar with the number of layers in the LSTM network. However, we note that with the same number of hidden layers, the training time for CNN is faster than LSTM. The reason for this is that CNN consists of a pooling layer in each convolutional layers which significantly reduce the amount of parameters and computation in the network. Therefore, for a fair comparison, we increased the complexity of CNN by increasing the number of convolutional layers (e.g., 4 layers) and the filter size for each layer so that the training time for both LSTM and CNN is approximately the same. It is important to point out that the high computational complexity of the proposed method is only during the training stage which is done offline. The trained network, however, is capable of offering a much faster solution during the real-time implementation.

TABLE I
VLC SIMULATION PARAMETERS

Parameter	Symbol	Value
Room dimension	-	$5 \times 5 \times 3$
LED half-intensity angle	$\Phi_{1/2}$	$60°$
Receiver FOV	Ψ_c	$85°$
PD physical area	A	1 cm^2
Optical filter gain	g_f	1
Optical concentrator gain	$g(\psi)$	1
PD responsivity	R	1 A/W
Reflection coefficient	ρ	0.8
Transmitted optical power	P_t	1 W
Number of subcarriers	k	128

IV. SIMULATION RESULTS

A. Effect of Pilot Numbers

We compare the performance of our LSTM-based signal detection method with the traditional LS estimation, and with CNN in terms of SERs for different SNRs. The simulation parameters are listed in Table I. From the results depicted in Fig. 2 and Fig. 3, it is clear that the proposed model is proven to always achieve better performance than both LS and CNN and can perform almost as good as ML detection in both fixed and random scenarios. LS is a linear estimator which may perform well in ideal scenarios but may give poor performance when a realistic environment is considered. In contrast, our DL model offers excellent performance since it can adapt to specific geometrical configurations and user behavior effects. Hence, focusing on the fixed scenario, a significant gain can be achieved in comparison to LS especially when fewer pilots were used. In Fig. 3, we can see that at SER $= 3.8e10^{-3}$, an SNR gain of more than 10 dB was obtained for LSTM-based approach compared to LS. This can be expected as even when the number of pilots is reduced, the DL model can still perform well since it has the ability to use the whole sequence of historical data to learn the channel and the user behavior and make reliable predictions of the transmitted signal. In comparison, LS gives poor performance due to the fact that the reduced number of pilots is not enough for LS to accurately estimate the channel. Therefore, it appears that even when the system have a very limited amount of channel information, the proposed model can still give excellent detection performance, very close to ML with full CSI.

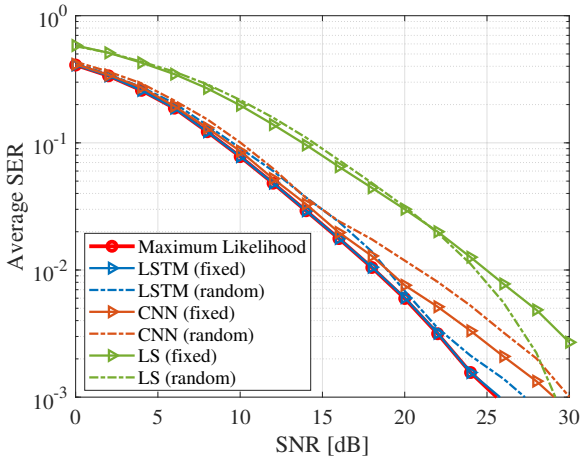


Fig. 2. Average SER versus SNR with pilot ratio of 1/8.

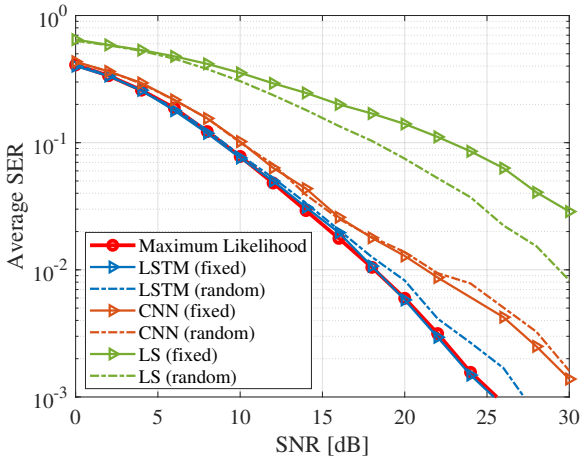


Fig. 3. Average SER versus SNR with pilot ratio of 1/32.

B. Hotspot Model

We have also simulated the signal detection problem using the proposed method which was trained based on a hotspot model. The hotspot model considers an attraction area in the room where the probability of the user to be within the area is higher than being elsewhere. We collect the training data in a similar manner as the previous approach for all the possible locations of the user. We then consider 3 types of scenarios where the probability of the user to be inside the hot spot area is 100%, 90% or 80%. To realize this hot spot model, for each scenario, we train the LSTM network by using 100%, 90% and 80% of the total data collected from all possible locations of the user inside the hot spot area, respectively. For the second and third case, the balancing 10% and 20% of the data were taken from the random user positions located outside the hot spot area. We then compare the performance between LSTM and LS in the case where the pilot ratio is 1/32. Fig. 4 presents the SER curves for the three different scenarios. As expected, the best performance similar to ML can be achieved by having 100% knowledge of the locations of where the user might be. Although the model would yield to a higher SER when the amount of training data based on the hot

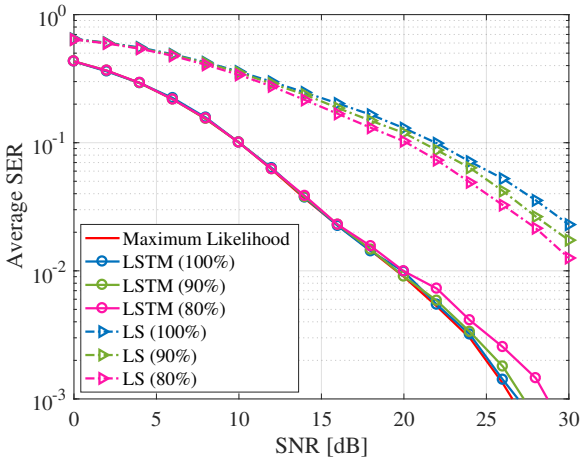


Fig. 4. Average SER versus SNR with different hotspot models.

spot area is reduced, the model still provide good performance in comparison to LS. Hence, from the result shown in Fig. 4, we can demonstrate that the deep learning model is also robust when applied for different user behaviors. It is important to note that as long as the model obtain sufficient training data based on user behaviors (e.g., the location within the room where the user would most probably be), it could significantly help with the learning process of the LSTM network and improve the detection performance of the system.

C. Size of Training Data

The size of the training data set depends on the complexity of the system as well as the complexity of the learning algorithms. Using too little training data may cause poor performance since the model will not be able to fully learn the diverse characteristics of the environment under study. Having too large training data may result in overfitting, in which the model corresponds too closely to the training data but fail to give good performance when new data is presented during the testing stage. Thus, we conduct several experiments to determine the suitable training data size and see which of them could provide the best SER vs SNR performance. Focusing on the fixed scenario, Table II and Table III show the effect of increasing the total number of training data for LSTM and CNN when the pilot ratio used is 1/8 and 1/32, respectively. From these two tables, we can see that increasing the length of training data for both deep learning methods noticeably helps to reduce the SNR penalties against ML when read at SER $= 3.8 \times 10^{-3}$. LSTM still gives better performance compared to CNN with increasing size of training data.

D. Complexity Analysis

It has been mentioned in previous works that LSTM algorithm is very efficient and is local in space and time [21]. This means that the storage requirements of the network does not depend on the input sequence length and at each time step, the computational complexity of an LSTM layer per weight is $\mathcal{O}(1)$. Hence, the total complexity of an LSTM at each time step is $\mathcal{O}(w)$ where w is the number of weights.

TABLE II
EFFECT OF TOTAL NUMBER OF TRAINING DATA FOR PILOT RATIO OF $1/8$

SNR penalty against ML at SER $= 3.8e10^{-3}$, (dB)			
Size of training data	2.7 x 10^6	1.4 x 10^7	2.7 x 10^7
LSTM	>10	0.9	0
CNN	>10	5.9	1.8

TABLE III
EFFECT OF TOTAL NUMBER OF TRAINING DATA FOR PILOT RATIO OF $1/32$

SNR penalty against ML at SER $= 3.8e10^{-3}$, (dB)			
Size of training data	2.7 x 10^6	1.4 x 10^7	2.7 x 10^7
LSTM	>10	5	2.1
CNN	>10	>10	5

Therefore, the time complexity for our model is $\mathcal{O}(\sum_{l=1}^{d} w_l)$ where l and d are the index and the number of LSTM layers, respectively while w_l denotes the number of weights for the l-th layer. In our simulation, the LSTM model consist of 3 layers with 50, 25 and 10 hidden units, respectively. The time complexity for all convolutional layers [24] for the CNN model is $\mathcal{O}(\sum_{l=1}^{d} n_{l-1} \cdot s_l^2 \cdot n_l \cdot m_l^2)$ where l and d are the index and the number of convolutional layers, respectively. We used 4 layers with the number of filters denoted by n_l for each layer is 8, 16, 32 and 64, respectively. $s_l = 3$ is the spatial size of the filter. Since there are 128 symbols and 26 SNR values, we have a total of 3328 data symbols for each subset of training data. We arrange the input sequence into a two dimensional (2-D) matrix to obtain an input data size of 58x58. Hence, the spatial size of the output feature map, m_l is 58. Lastly, since our input data is 2-D, the number of input channel at the l-th layer, denoted by n_{l-1} is 1. Note that the size of the training data exceeds the number of parameters to be learned by both of the neural network models. It was mentioned earlier that channel estimation using LS requires low complexity and can be obtained by a simple division of the received pilot symbols under the effect of the optical channel over the transmitted pilot symbols. However channel estimation with LS gives insufficient performance as opposed to the DL techniques.

V. CONCLUSION

Our results confirm that the proposed DL approach has the ability to learn and analyze the complicated characteristics of the optical channel and the user behavior and is able to provide a good generalization ability when operating online. It is observed that the proposed DL method, only with limited CSI, can provide a close performance to the optimal ML detection scheme which requires perfect CSI. Furthermore, the robust performance of the proposed DL scheme in different environments is shown. Future works for this study will include the application of deep learning to a multi-user scenario in the presence of the effect of link blockage, mobility, etc.

REFERENCES

[1] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using led lights," *IEEE Trans. Consum. Electron.*, vol. 50, pp. 100–107, Feb 2004.

[2] S. Wu, H. Wang, and C. Youn, "Visible light communications for 5g wireless networking systems: from fixed to mobile communications," *IEEE Network*, vol. 28, pp. 41–45, Nov 2014.

[3] M. D. Soltani, X. Wu, M. Safari, and H. Haas, "Bidirectional user throughput maximization based on feedback reduction in lifi networks," *IEEE Trans. Commun.*, vol. 66, pp. 3172–3186, July 2018.

[4] P. G. Pachpande, M. H. Khadr, A. F. Hussein, and H. Elgala, "Visible light communication using deep learning techniques," in *2018 IEEE 39th Sarnoff Symposium*, pp. 1–6, Sep. 2018.

[5] H. Lee, I. Lee, T. Q. S. Quek, and S. H. Lee, "Binary signaling design for visible light communication: a deep learning framework," *Opt. Express*, vol. 26, pp. 18131–18142, Jul 2018.

[6] H. Lee, I. Lee, and S. Lee, "Deep learning based transceiver design for multi-colored vlc systems," *Optics Express*, vol. 26, p. 6222, 03 2018.

[7] H. Ye, G. Y. Li, and B. Juang, "Power of deep learning for channel estimation and signal detection in ofdm systems," *IEEE Wireless Commun. Lett.*, vol. 7, pp. 114–117, Feb 2018.

[8] J. Wang, C. Jiang, H. Zhang, X. Zhang, V. C. M. Leung, and L. Hanzo, "Learning-aided network association for hybrid indoor lifi-wifi systems," *IEEE Trans. Veh. Technol.*, vol. 67, pp. 3561–3574, April 2018.

[9] S. Ma, J. Dai, S. Lu, H. Li, H. Zhang, C. Du, and S. Li, "Signal demodulation with machine learning methods for physical layer visible light communications: Prototype platform, open dataset, and algorithms," *IEEE Access*, vol. 7, pp. 30588–30598, 2019.

[10] B. Karanov, G. Liga, V. Aref, D. Lavery, P. Bayvel, and L. Schmalen, "Deep learning for communication over dispersive nonlinear channels: Performance and comparison with classical digital signal processing." [online]Available:https://arxiv.org/abs/1910.01028, 2019.

[11] M. Zhang, Y. Zeng, Z. Han, and Y. Gong, "Automatic modulation recognition using deep learning architectures," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1–5, June 2018.

[12] N. Farsad and A. J. Goldsmith, "Detection algorithms for communication systems using deep learning," *CoRR*, vol. abs/1705.08044, 2017.

[13] X. Ouyang, C. Zhang, P. Zhou, and H. Jiang, "Deepspace: An online deep learning framework for mobile big data to understand human mobility patterns," *CoRR*, vol. abs/1610.07009, 2016.

[14] M. Ozturk, M. Gogate, O. Onireti, A. Adeel, A. Hussain, and M. A. Imran, "A novel deep learning driven, low-cost mobility prediction approach for 5g cellular networks: The case of the control/data separation architecture (cdsa)," *Neurocomputing*, vol. 358, pp. 479 – 489, 2019.

[15] Z. Zhang, Y. Zhu, W. Zhu, H. Chen, X. Hong, and J. Chen, "Iterative point-wise reinforcement learning for highly accurate indoor visible light positioning," *Optics Express*, vol. 27, p. 22161, Aug 2019.

[16] M. Hulea, Z. Ghassemlooy, and S. Rajbhandari, "A spiking neural network with visible light communications," in *2018 11th International Symposium on Communication Systems, Networks Digital Signal Processing (CSNDSP)*, pp. 1–5, July 2018.

[17] W. Zhang *et al.*, "Neural networks for papr reduction in optical ofdm signal transmission," in *2019 21st International Conference on Transparent Optical Networks (ICTON)*, pp. 1–4, July 2019.

[18] H. Schulze, "Frequency-domain simulation of the indoor wireless optical communication channel," *IEEE Trans. Commun.*, vol. 64, pp. 2551–2562, June 2016.

[19] M. D. Soltani, A. A. Purwita, Z. Zeng, H. Haas, and M. Safari, "Modeling the random orientation of mobile devices: Measurement, analysis and lifi use case," *IEEE Trans. Commun.*, vol. 67, pp. 2157–2172, March 2019.

[20] S. D. Dissanayake and J. Armstrong, "Comparison of aco-ofdm, dco-ofdm and ado-ofdm in im/dd systems," *J. Lightw. Technol.*, vol. 31, pp. 1063–1072, April 2013.

[21] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, pp. 1735–80, 12 1997.

[22] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Commun. Mag.*, vol. 57, pp. 114–119, June 2019.

[23] E. Tsironi, P. Barros, C. Weber, and S. Wermter, "An analysis of convolutional long short-term memory recurrent neural networks for gesture recognition," *Neurocomputing*, vol. 268, pp. 76 – 86, 2017.

[24] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5353–5360, June 2015.