

Deep Neural Network-based Channel Estimation In OFDM Systems

Rishika Chauhan, Shefali Sharma, Rahul Pachauri
Department of Electronics and Communication Engineering
Jaypee University of Engineering and Technology
Guna-473226, Madhya Pradesh, India

Abstract—In this research article, an attempt has been made to improve the performance of channel estimation in OFDM systems with the help of deep neural network (DNN). A bi-directional long short term memory (bi-LSTM) based DNN model is proposed and trained using three commonly used optimization algorithms, stochastic gradient descent with momentum (SGDm), root mean square with propagation (RMSprop) and adaptive moment estimation (Adam). Performance analysis and comparison of these algorithms has been carried out using least square (LS) and minimum mean square error (MMSE) estimators with different size of inputs. The findings have revealed that the proposed DNN model can be used as channel estimator in OFDM systems without requiring any prior information of channel statistics.

Keywords—Deep neural network; Channel estimation; OFDM; Optimization algorithms

Nomenclatures:

X	transmitted sequence vector
Y	received sequence vector
N	number of subcarriers
H	channel matrix
I	interference
W	additive white gaussian noise
R	correlation matrix
n	OFDM symbol no.
k	subcarrier no.
m	first moment vector
v	second moment vector
w	weight
b	bias
c	cell state
f	forget gate
o	output gate
i	input gate
h	final output of LSTM cell

Greek Letter

α	learning rate
β	moving average
π	pi

σ^2	variance
θ	network parameter

Abbreviations

Adam	adaptive moment estimation
ANN	artificial neural network
AWGN	additive white gaussian noise
BER	bit error rate
BPNN	back propagation neural network
BPSK	binary phase shift keying
CP	cyclic prefix
DNN	deep neural network
GA	genetic algorithm
ISI	intersymbol interference
LM	Levenberg Marquardt
LS	least square
LSTM	long short term memory
MMSE	minimum mean square error
MSE	mean square error
OFDM	orthogonal frequency division multiplexing
QAM	quadrature amplitude modulation
QPSK	quadrature phase shift keying
RMS	proproot mean square propagation
RNN	recurrent neural network
SER	symbol error rate
SGD	stochastic gradient descent
SNR	signal to noise ratio

I. INTRODUCTION

Orthogonal frequency division multiplexing (OFDM) is a well-known modulation technique adopted in modern wireless systems to assuage frequency selective fading in wireless channels as it has the ability to mitigate the intersymbol interference (ISI) produced by delay spread of wireless channels. Channel estimation is one of the major issues in OFDM system since the response of the channel vary rapidly with time due to the mobility of transmitter, receiver or scattering objects [1]. Lot of attempts have been

made by several researchers to estimate the effect of channel in OFDM system accurately. Conventional methods like LS and MMSE are mostly used for pilot assisted channel estimation [2]. In earlier works, it has been observed that LS estimator shows inadequate performance, however, it does not require any prior channel statistics while MMSE estimator provides better performance than LS but at the cost of higher complexity [3]–[5]. To reduce its complexity, several techniques are introduced in the literature [6].

Recently, artificial neural networks (ANNs) have drawn attention to estimate the channel with less complexity [7]. ANNs consist of several neurons that operate in parallel. The neurons are interconnected through weighted inputs and provides the ability of learning, recalling and generalizing the training data. Cui and Tellambura [8] have used radial basis networks (a type of neural network) for estimating the channel in OFDM systems. Backpropagation neural network (BPNN) is a multilayer neural network which is used as a channel estimator by Tasnibar et al. [9]. The authors have reported that the performance of MMSE estimator is better than that of LS and BPNN but with higher complexity. Further, to improve the channel estimation performance, genetic algorithm (GA) is combined with BPNN by Cheng et al and they have reported that GA based BPNN is superior than the conventional BPNN [10]. Some authors have also proposed deep learning applications for channel estimation [11]. Hao et al have proposed DL based structure for channel estimation and signal detection in OFDM systems and they have shown that deep learning models can work better than traditional methods with enough pilots [12]. In [13], DL-based channel estimation network (CENet) and channel conditioned recovery net (CCRNet) are employed for joint channel estimation and signal detection in OFDM systems. The authors have demonstrated that both proposals provide good generalization ability and robustness toward the channel parameter variation.

In this research study, a deep neural network (DNN) model is proposed for estimating the channel and its performance is compared with the conventional channel estimators. Three different optimization algorithms have been used for training of proposed DNN to obtain the efficient estimator with lower symbol error rate (SER). The proposed DNN is trained offline due to the large number of network parameters required for updation. Then the trained DNN is employed online so that the required information can be recovered. The simulation results have shown that proposed DNN outperforms both LS and MMSE estimator with limited number of subcarriers and pilots. Also, the results have demonstrated that DNN based estimator using Adam optimizer can be efficiently used in OFDM systems to boost the transmission rate of the system.

II. SYSTEM MODEL AND METHODOLOGY USED

A. System Architecture

The block diagram of an OFDM system with DNN based channel estimation is shown in Fig. 1.

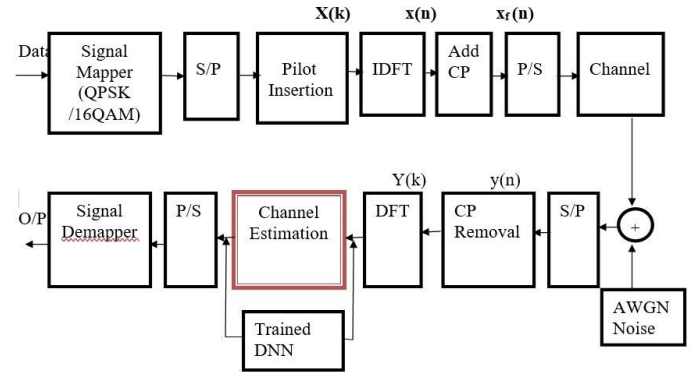


Fig. 1: DNN based OFDM system architecture

A binary data stream is first modulated using common modulation technique such as Quadrature Phase Shift Keying (QPSK) or Quadrature Amplitude Modulation (QAM). These data symbols are converted into parallel streams. After insertion of the pilot symbols, the data sequence $X(k)$ is transformed into time domain signal $x(n)$ via IDFT block [14], i.e.

$$x(n) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X(k) e^{i2\pi n k / N} \quad 0 \leq n \leq N-1 \quad (1)$$

Cyclic prefix is used for mitigation of inter symbol interference (ISI). Then, the data sequence passes through the channel with impulse response $h(n)$ of length L .

$$y_f(n) = x_f(n)h(n) + w(n) \quad (2)$$

where, $w(n)$ is the additive white gaussian noise (AWGN) which gets added in the signal during transmission through the channel. At the receiver, after removal of the cyclic prefix, DFT block is used to transform signal back to the frequency domain.

$$y(n) = CPremove\{y_f(n)\} \quad n = 0, 1, \dots, N-1 \quad (3)$$

$$Y(k) = DFT\{y(n)\} \quad k = 0, 1, \dots, N-1 \quad (4)$$

Then, pilot signals are extracted and the estimated channel response $H_{est}(k)$ is obtained for the data sequence in the channel estimation block.

$$Y(k) = X(k)H(k) + I(k) + W(k) \quad (5)$$

$$X_{est}(k) = \frac{Y(k)}{H_{est}(k)} \quad (6)$$

Once the transmitted signal is estimated, then it is demodulated to get the desired information.

B. LS and MMSE Estimator

Least Square (LS) Estimation

The LS estimator is used to minimize the squared error between the received and original signal without needing

any prior knowledge of channel statistics. The LS estimate of the channel is given by:

$$\hat{H}_{LS} = (X^H X)^{-1} X^H Y = X^{-1} Y \quad (7)$$

where, X is the input sequence vector and Y is the received sequence vector.

Minimum mean square error (MMSE) Estimation

The MMSE estimator is used to minimize the mean square error (MSE) by exploiting the second order statistics of the channel. The MMSE estimate of the channel is defined by:

$$\hat{H}_{MMSE} = R_{H\tilde{H}} \left(R_{H\tilde{H}} + \frac{\sigma_z^2}{\sigma_x^2} I \right)^{-1} \tilde{H} \quad (8)$$

where, H represents the actual channel response.

C. Deep Neural Network (DNN) Architecture

The structure of proposed deep neural network (DNN) with five significant layers is presented in this section. The neural network with large number of layers is termed as deep neural network (DNN). The proposed DNN for estimating the channel consists of sequence input layer followed by bi-LSTM, fully connected, SoftMax and classification layer. The bi-LSTM layer is formed using two independent recurrent neural networks (RNNs) that can learn long-standing associations between the time steps of data sequence. In this layer, 20 hidden units are used. Input size is varied in accordance with the number of subcarriers and fully connected layer of four classes is included.

As bi-LSTM layer is composed of bidirectional RNNs which allow the network to have both forward and backward knowledge of the sequence at each time step. It provides the opportunity to save the information from both past and future. Fig. 2. represents the structure of the bidirectional LSTM layer.

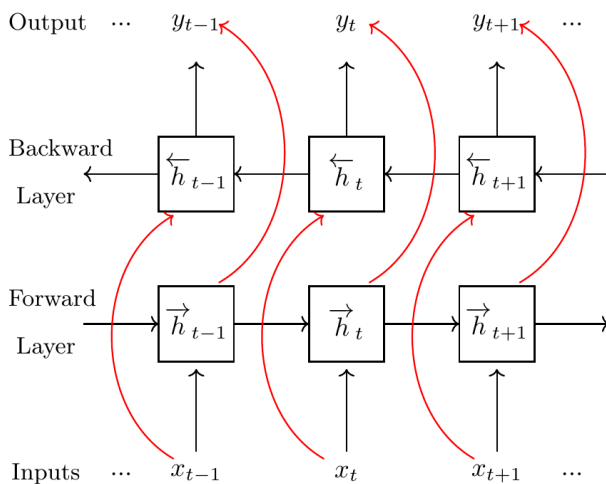


Fig. 2: Structure of bidirectional LSTM

A single cell of an LSTM network is shown in Fig. 3 which consists of an input, output and forget gate. The following relations are used to implement the LSTM cell [15]:

$$f_t = \sigma(W_{fh}h_{t-1} + W_{fx}x_t + b_f) \quad (9)$$

$$i_t = \sigma(W_{ih}h_{t-1} + W_{ix}x_t + b_i) \quad (10)$$

$$\tilde{c}_t = \tanh(W_{\tilde{c}h}h_{t-1} + W_{\tilde{c}x}x_t + b_{\tilde{c}}) \quad (11)$$

$$c_t = f_t \Delta c_{t-1} + i_t \Delta \tilde{c}_t \quad (12)$$

$$o_t = \sigma(W_{oh}h_{t-1} + W_{ox}x_t + b_o) \quad (13)$$

$$h_t = o_t \tanh(c_t) \quad (14)$$

where, W_{fh} , W_{fx} , W_{ih} , W_{ix} , $W_{\tilde{c}h}$, $W_{\tilde{c}x}$, W_{oh} , W_{ox} and b_f , b_i , $b_{\tilde{c}}$, b_o denote the weights and biases respectively. σ is the sigmoid function. f_t , c_t , i_t and o_t represent the forget gate, cell state, input gate and output gate.

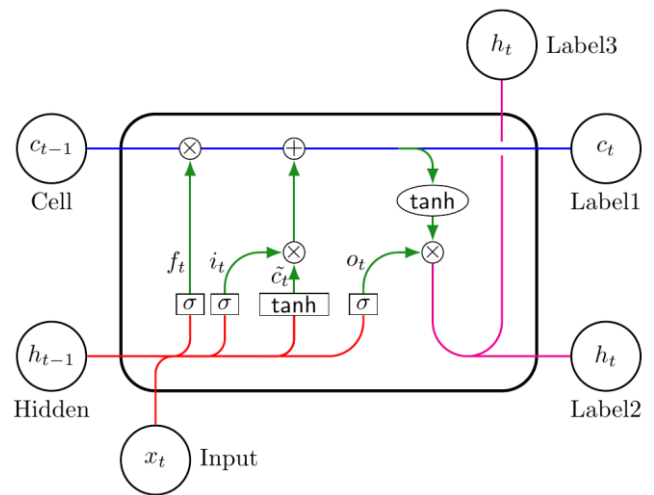


Fig. 3: Single LSTM cell

III. DNN BASED CHANNEL ESTIMATION

The proposed DNN is trained offline which takes the received data as its input and produces the transmitted data at its output. The channel model used is the narrowband Rayleigh fading channel. The training data is generated for a single user OFDM system in which the OFDM frame consists of pilot and transmitted symbols that are randomly generated. The received OFDM frame is recovered and considered as input to the DNN model. To minimize the error between the trained output and the original transmitted data, the proposed DNN model is trained using three different optimization algorithms. The performance of the proposed estimator is analyzed using these three optimizers and compared with the traditional methods of channel estimation. The following three optimizers used are stochastic gradient descent with momentum (SGDm), root mean square propagation (RMSprop) and adaptive moment estimation (Adam).

A. Stochastic gradient descent (SGD)

This is the common optimization algorithm used for faster convergence of neural networks. It exploits few samples that are stochastically selected from the whole dataset to perform every iteration. SGD is computationally less expensive even it requires more iterations to reach the global minimum than gradient descent [16], [17]. In this article, SGD with momentum (SGDm) is used to improve the convergence speed.

In SGDm, a moving average of gradient is computed to update the network parameters. The first moment vector (m) and the network parameters (θ) at iteration 't' are updated as follows [16], [17]:

$$m_t = \beta m_{t-1} + (1 - \beta) \text{grad}(\theta_{t-1}) \quad (15)$$

$$\theta_t = \theta_{t-1} - \alpha m \quad (16)$$

where the term β controls the moving average. The default value of β is 0.9. α is the learning rate and θ denotes the network parameters (weights and biases) to be updated.

B. Root mean square propagation (RMSprop)

This algorithm is also based on gradient descent algorithm. RMSprop considers the moving average of the squares of the recent gradients instead of using all. It has the ability to reduce the loss function continuously to reach the minimum throughout the training process.

In RMSprop, the second moment vector (v) and the network parameters are updated using the following relations [16], [17]

$$v_t = \beta v_{t-1} + (1 - \beta) \text{grad}(\theta_{t-1})^2 \quad (17)$$

$$\theta_t = \theta_{t-1} - \alpha \left(\frac{\text{grad}(\theta_{t-1})}{\sqrt{v_t + \epsilon}} \right) \quad (18)$$

where, the term ϵ is very small and used for numerical stability.

C. Adaptive moment estimation (Adam)

This algorithm is very popular as it combines the benefits of RMSprop and momentum. This provides faster optimization as it uses adaptive learning rates to update the network parameters and is preferred for training of deep neural networks. Using Adam optimizer, the network parameters are updated as follows [16], [17]:

$$m_t = \beta m_{t-1} + (1 - \beta) \text{grad}(\theta_{t-1}) \quad (19)$$

$$v_t = \beta v_{t-1} + (1 - \beta) \text{grad}(\theta_{t-1})^2 \quad (20)$$

$$\hat{m}_t = \frac{m_t}{(1 - \beta_1^t)} \quad (21)$$

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} \quad (22)$$

$$\theta_t = \theta_{t-1} - \alpha \left(\frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \right) \quad (23)$$

where, β_1 and β_2 denote the exponentially weighted averages (moving averages) for SGDm and RMSprop respectively. \hat{m}_t and \hat{v}_t are the bias corrected values of corresponding m_t and v_t .

IV. OBSERVATIONS, RESULTS AND DISCUSSIONS

In this research study, the proposed DNN is trained according to the generated data sets and used for estimating the channel. For different signal to noise ratios (SNRs), symbol error rates (SERs) obtained from conventional methods like LS and MMSE, and proposed DNN are compared for different optimization algorithms.

The dataset for training and validation is generated for a single subcarrier. The received OFDM packet consist of data symbols that are interleaved with the pilot symbols. Table I depicts the simulation parameters of OFDM system whereas, the training parameters for the proposed DNN are shown in Table II.

TABLE I: OFDM simulation parameters

Parameter	Particular
Number of subcarriers	64, 256
Modulation type	QPSK
Guard interval type	Cyclic prefix (CP)
Length of pilot sequence	16
Noise model	Additive white Gaussian noise (AWGN)
Channel model	Rayleigh fading channel
Number of transmitted symbol	10000

TABLE II: Proposed DNN training parameters

Parameter	Particular
Input size	256, 1024
Fully connected layers	4
BiLSTM layer size	20 hidden neurons
Minibatch Size	300
Number of epochs	1000
Loss function	crossentropy
Optimizers	Adam, RMSprop, SGDm

A comparative analysis is done for evaluating the performance of three estimators at different subcarriers. When only 64 subcarriers are used, the proposed DNN outperforms the conventional methods at all signal to noise ratios as shown in Fig. 4, 5 and 6. If the performance of three

optimization algorithms is considered, then it can be concluded that Adam optimizer performs very well for the developed model and employed dataset.

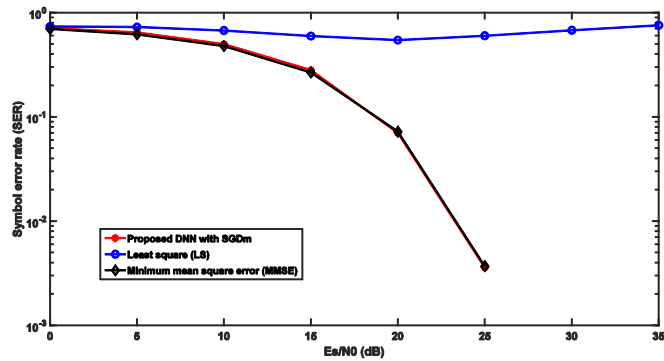


Fig. 4: SER performance for 64 subcarriers with SGDM optimizer

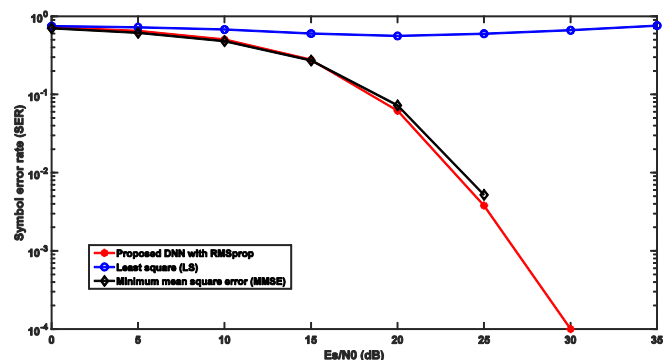


Fig. 5: SER performance for 64 subcarriers with RMSprop optimizer

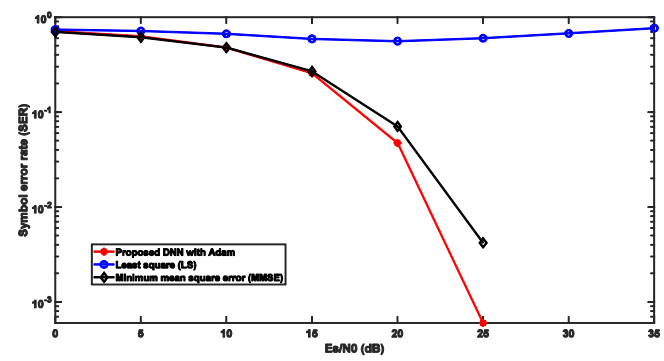


Fig. 6: SER performance for 64 subcarriers with Adam optimizer

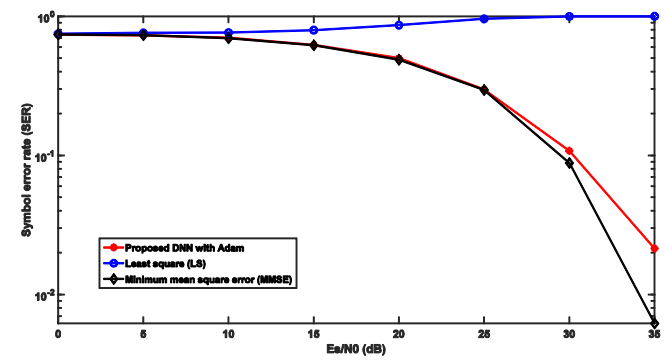


Fig. 7: SER performance for 256 subcarriers with Adam optimizer

It can be observed from Fig. 7, 8, and 9, as the number of subcarriers gets increased to 256, the proposed estimator

shows the better performance than LS estimator and comparable performance to the MMSE estimator at lower SNR values. The LS estimator has the poor performance as it requires no prior channel information but the MMSE estimator has excellent performance at more subcarriers as it involves the second order statistics of the channel. Also, the developed DNN model is able to reduce SER with growing SNR without requiring any prior information of channel statistics, which makes it propitious for channel estimation in OFDM systems.

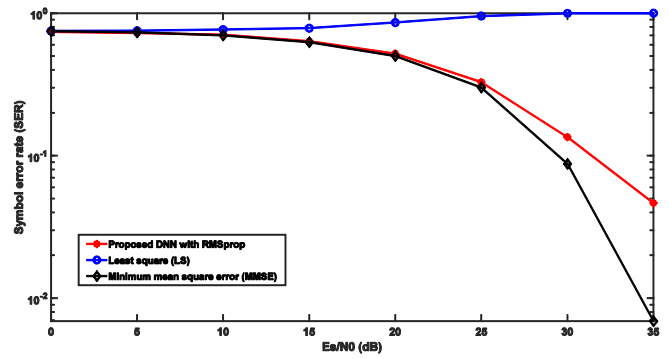


Fig. 8: SER performance for 256 subcarriers with RMSprop optimizer

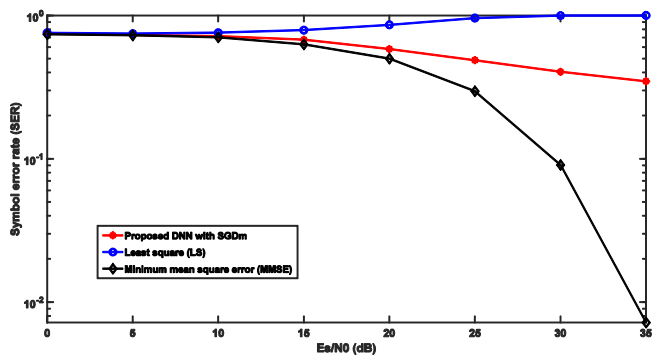


Fig. 9: SER performance for 256 subcarriers with SGDM optimizer

A. Performance of Optimization algorithms

Optimization algorithms, also known as optimizers, play a crucial role to train and improve the performance of deep neural networks (DNNs). The performance of these optimizers can be analyzed in terms of developed model and generated dataset to obtain the efficient estimator. The SER performance of commonly used optimizers viz. SGDM, RMSprop and Adam on the proposed DNN is shown in Fig. 10 and 11. It can be easily observed from these figures that Adam optimizer outperforms RMSprop and SGDM optimizers. Depending on the number of subcarriers, the same optimization algorithms differ in performance and the SGDM optimizer shows the inferior performance. It is observed that the proposed DNN trained with Adam optimizer gives the excellent performance among the three optimization algorithms as shown in Table III. Therefore, the combination of DNN and Adam optimizer can be preferred in OFDM communication systems for efficient channel estimation.

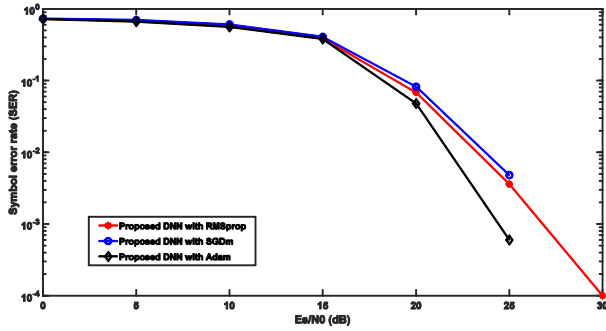


Fig. 10: Performance comparison of three optimizers for 64 subcarriers

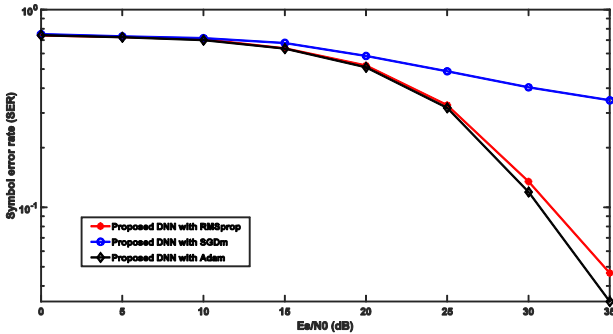


Fig. 11: Performance comparison of three optimizers for 256 subcarriers

TABLE III: SER performance for 64 subcarriers (256 input size)

Optimizers	SER	
	SNR(20 dB)	SNR (25 dB)
Adam	0.4	0.006
RMSprop	0.6	0.06
SGDm	0.7	0.05

V. CONCLUSIONS

The current study addresses the development of a deep neural network (DNN) using bidirectional LSTM neural network to improve the performance of channel estimation in OFDM systems. Different learning algorithms viz. SGDM, RMSprop and Adam have been used for optimization of proposed DNN. The effectiveness of the proposed DNN model is investigated for different input size (or different number of subcarriers) and a comparative analysis is performed with the traditional methods like LS and MMSE estimators using three optimizers. The obtained results reveal that the performance of LS and MMSE estimator lacks the performance of proposed estimator for 64 subcarriers (256 input size). Out of three optimizers, Adam optimizer shows the excellent performance as it achieves the SER value of 0.006 at 25dB (SNR) while it is 0.06 and 0.05 for RMSprop

and SGDM optimizers respectively, therefore, it can be concluded that the proposed DNN model trained with Adam optimizer can be efficiently used as a channel estimator in OFDM communication systems.

REFERENCES

- [1] T. Hwang, C. Yang, S. Member, G. Wu, S. Li, and G. Y. Li, "OFDM and Its Wireless Applications: A Survey," vol. 58, no. 4, pp. 1673–1694, 2009.
- [2] M. Ozdemir and H. Arslan, "Channel Estimation for Wireless OFDM systems," *IEEE Commun. Surv. Tutorials*, vol. 9, no. 2, pp. 18–48, 2007.
- [3] K. Liu and K. Xing, "Research of MMSE and LS channel estimation in OFDM systems," *2nd Int. Conf. Inf. Sci. Eng. ICISE2010 - Proc.*, vol. 1, no. 1, pp. 2308–2311, 2010.
- [4] S. Ahmed Ghauri and M. Farhan Sohail, "Implementation of Ofdm and Channel Estimation Using Ls and Mmse Estimators," *Int. J. Comput. Electron. Res.*, vol. 2, no. 1, pp. 41–46, 2013.
- [5] A. Sahu and A. Khare, "A Comparative Analysis of LS and MMSE Channel Estimation Techniques for MIMO-OFDM System," *J. Eng. Res. Appl. www.ijera.com*, vol. 4, no. 6, pp. 162–167, 2014.
- [6] M. B. Sutar and V. S. Patil, "Complexity Reduction Techniques for MMSE Channel Estimator in OFDM," *IEEE Veh. Technol. Conf.*, vol. 08, no. 6, pp. 18–24, 2018.
- [7] C. H. Cheng, Y. H. Huang, and H. C. Chen, "Enhanced channel estimation in OFDM systems with neural network technologies," *Soft Comput.*, vol. 23, no. 13, pp. 5185–5197, 2019.
- [8] T. Cui and C. Tellambura, "Channel estimation for OFDM systems based on adaptive radial basis function networks," *IEEE Veh. Technol. Conf.*, vol. 60, no. 1, pp. 608–611, 2004.
- [9] N. Taşpinar and M. N. Seyman, "Back propagation neural network approach for channel estimation in OFDM system," *Proc. - 2010 IEEE Int. Conf. Wirel. Commun. Netw. Inf. Secur. WCNIS 2010*, pp. 265–268, 2010.
- [10] C. C. Y. H. H. Chen, "Channel estimation in OFDM systems using neural network technology combined with a genetic algorithm," *Soft Comput.*, vol. 20, no. 10, pp. 4139–4148, 2016.
- [11] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, "Deep Learning-Based Channel Estimation," *IEEE Commun. Lett.*, vol. 23, no. 4, pp. 652–655, 2019.
- [12] H. Ye, G. Y. Li, and B. H. Juang, "Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems," *IEEE Wirel. Commun. Lett.*, vol. 7, no. 1, pp. 114–117, 2018.
- [13] X. Yi and C. Zhong, "Deep learning for joint channel estimation and signal detection in OFDM systems," *arXiv*, pp. 1–5, 2020.
- [14] R. T. Sataloff, M. M. Johns, and K. M. Kost, *Orthogonal Frequency Division Multiplexing for Wireless Communications*. .
- [15] T. Faghani, A. Shojaeifard, K. K. Wong, and A. H. Aghvami, "Recurrent neural network channel estimation using measured massive MIMO data," *IEEE Int. Symp. Pers. Indoor Mob. Radio Commun. PIMRC*, vol. 2020-Augus, 2020.
- [16] J. Zhang, "Gradient Descent based Optimization Algorithms for Deep Learning Models Training," *arXiv*, 2019.
- [17] D. Soydaner, "A Comparison of Optimization Algorithms for Deep Learning," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 34, no. 13, pp. 1–26, 2020.