Journal of Electrical Engineering, Vol. 74, No. 3, 2023, pp. 167–176

sciendo

PAPER

Effective deep learning-based channel state estimation and signal detection for OFDM wireless systems

Hassan A. Hassan[1,*], Mohamed A. Mohamed [1], Mohamed H. Essai[1], Hamada Esmaiel[2], Ahmed S. Mubarak [2] and Osama A. Omer[2]

Deep learning (DL) algorithms can enhance wireless communication system efficiency and address numerous physical layer challenges. Channel state estimation (CSE) and signal detection (SD) are essential parts of improving the performance of an OFDM wireless system. In this context, we introduce a DL model as an effective alternative for implicit CSE and SD over Rayleigh fading channels in the OFDM wireless system. The DL model is based on the gated recurrent unit (GRU) neural network. The proposed DL GRU model is trained offline using the received OFDM signals related to the transmitted data symbols and added pilot symbols as inputs. Then, it is implemented online to accurately and directly detect the transmitted data. The experimental results using the metric parameter of symbol error rate show that, the proposed DL GRU-based CSE/SD provides superior performance compared with the traditional least square and minimum mean square error estimation methods. Also, the trained DL GRU model exceeds the existing DL channel estimators. Moreover, it provides the highest CSE/SD quality with fewer pilots, short/null cyclic prefixes, and without prior knowledge of the channel statistics. As a result, the proposed DL GRU model is a promising solution for CSE/SD in OFDM wireless communication systems.

Keywords: OFDM, channel state estimation, signal detection, deep learning, GRU

1. Introduction

Orthogonal frequency-division multiplexing (OFDM) is a prominent modulation technology utilized in current wireless communication systems. Because of its superior bit error rate (BER) quality, high spectrum efficacy, significant resistance to multi-path fading, and strong resistance to interference, OFDM is a strong choice for high data rates in next-generation wireless communication systems [1].

The performance of an OFDM system is highly dependent on the channel estimation and signal detection techniques used. In general, the influence of channels in communication systems distorts the received signal. The channel influence at the receiver, channel state information (CSI), must be precisely estimated to retrieve the transferred symbols. In general, the receiver estimates the CSI by employing pilot symbols that are known to both the receiver and the transmitter [2, 3].

There are various conventional approaches employed in OFDM systems for channel estimation, including least square (LS) and minimum mean square error (MMSE). The LS estimator is constructed with minimal complexity without requiring any knowledge of channel statistics. But, it ignores noise interference in the computation operation, which leads to inadequate performance in a complicated communication context [4-6]. In contrast, the MMSE estimator considers the effect of noise to enhance the accuracy of the channel estimate. It is, however, more complicated than the LS estimation method since it requires prior knowledge of the statistical characteristics of the channel. In several situations, such statistical data is either impossible to gather or varies rapidly in a short period of time [7].

In addition to the traditional model based on channel estimation and signal detection techniques, new machine learning (ML)-based models that use deep learning algorithms to perform channel estimation and signal detection have recently emerged as an efficient alternative [8-12]. DL has recently received a lot of interest in wireless communications systems. Several methods have been developed in DL-based communication systems to improve the performance of various present approaches, such as channel estimation, channel equalization, signal decoding, and radio resource allocation [9-24].

In terms of the channel estimate and signal detection applications, the authors in [13], introduce a new channel estimation framework for the MIMO-OFDM system aided by several deep neural network (DNN) structures. The proposed DL-based estimators outperformed the traditional LS and LMMSE estimations. In [14], the authors presented a DL-based channel estimation approach for the IEEE 802.11p standard. The experimental results indicated that the proposed channel estimating technique outperformed conventional channel estimators significantly. The authors in [15], suggested two DNN models for channel estimation with the aid of a pilot signal in underwater acoustic OFDM system. The numerical results showed that the proposed models achieved superior performance compared to the backpropagation neural network

[1] Department of Electrical Engineering, Faculty of Engineering, Al-Azhar University, Qena 83513, Egypt
[2]Aswan Faculty of Engineering, Aswan University, Abulrish 81542, Egypt
hassanali2720.el@azhar.edu.eg

https://doi.org/10.2478/jee-2023-0022, Print (till 2015) ISSN 1335-3632, On-line ISSN 1339-309X

(BPNN) model and the traditional MMSE and LS estimators. In [16], the author proposed a DL LSTM model for channel estimation and signal detection with the aid of pilot signals in a multipath OFDM system. The experimental results showed the effectiveness of the proposed model compared with the MMSE and LS estimations in a limited number of pilots. A deep residual channel estimation network (ReEsNet) was proposed for channel estimation in OFDM system [17]. The proposed model performed better with minimum complexity requirements compared to the LMMSE estimator. A deep complex-valued convolutional network (DCCN) was proposed to retrieve information from OFDM time-domain signals [18]. The simulation results demonstrated the proposed model outperformed the traditional estimation methods in Rayleigh fading channels with mobility and different delay spreads. In [19], the authors suggested a DL-based channel estimation network named, ChanEstNet in high-speed scenarios for the OFDM systems. The suggested model employed the integration of LSTM and CNN. Moreover, it has lower computing complexity and much superior performance compared to conventional approaches. In [20], an online signal detector model for OFDM wireless communication system was developed by using LSTM NN. The experimental results indicated that the employed model has a lower BER than the traditional algorithms. Using the assistance of a CNN and batch normalization layer, the authors in [21] suggested a recurrent neural network (RNN) with bidirectional long short-term memory (BiLSTM) structure for SD in a time-varying OFDM system. The authors in [22] introduce a system that combines compressive sensing (CS) and DL BiLSTM structure to perform joint CSE/SD in a MIMO-OFDM system.

RNNs are designed to process sequence data and have demonstrated outstanding performance in numerous time series tasks, especially those that involve short sequences. RNN can therefore be used as a CSE to improve estimation and detection performance. Although RNNs combined with BiLSTM architecture outperform standard RNNs in terms of capturing long-term dependencies, the architecture of combining RNNs with BiLSTM is rather complicated.

Regarding the above challenges, in this study, we adopt a DL architecture operating in an end-to-end manner that integrates the functions of CSE and SD for OFDM wireless communication systems over Rayleigh fading channels. The proposed model is based on the gated recurrent unit (GRU) deep learning NN, which is an efficient form of RNNs. The DL GRU model provides an accurate CSI, correctly retrieves transmitted data, and reduces the OFDM system's receiver architecture. Below is a summary of the main contributions of this work:

- Use the GRU network model to build a light computational DNN method for joint CSE and SD at the OFDM receiver. As a result, the proposed DL GRU structure offers an efficient solution to reduce the spectrum of resources necessary for CSE and SD in OFDM wireless systems.
- The proposed DL GRU framework is first trained offline using the simulation data set findings. As a result, the DL model can estimate the channel information implicitly. The trained DL model is then used online to directly predict/ retrieve the transmitted data without explicitly CSI estimating.
- We demonstrate the efficiency of the proposed DL GRU-based CSE/SD framework by comparing it to traditional LS and MMSE estimation techniques, according to the SER vs. SNR criterion. Furthermore, the suggested framework's performance is compared to data-driven approaches such as the DL BiLSTM model used in [21, 22].
- The performance of the examined estimators is tested in several simulated situations with cyclic prefix lengths, which are not discussed in [21, 22], and variable pilot density. Also, there is no prior channel statistics knowledge present.
- The proposed DL GRU model is trained by the adaptive moment estimation (Adam) optimizer in all simulation scenarios. In addition, three various loss functions to get the most effective DL GRU model.

According to the simulation results, the proposed DL GRU model achieves superior SER performance compared to the conventional LS and MMSE estimation methods. Furthermore, it outperforms the DL BiLSTM model used in [21, 22] in terms of robustness when the CP is omitted, limited training pilots are utilized, and no prior channel statistics knowledge is present. On the other hand, under the limited number of pilots and short/null CP, the proposed DL GRU with the mean absolute error (MAE) loss function achieves the lowest SER performance. In contrast, the proposed DL GRU with the conventional "cross-entropy" loss function outperforms the model with the sum of squared error (SSE) loss function regarding SER performance.

The rest of this paper is organized as follows. Section 2 discusses the OFDM system model. Section 3 describes the proposed DL GRU model based on CSE and SD and the model training process. In Section 4, the performance of the proposed DL model is examined using simulation results under different scenarios. Finally, the study is concluded in Section 5.

2. System architecture

Figure 1 illustrates a simplified model of the employed OFDM system. An OFDM system with a single user for the current study was employed. The transmitting and receiving elements of the OFDM system represented in Fig. 1 are identical to those used in traditional systems [24].

On the transmitting side, the input binary data is first generated and mapped to data symbols using a specific modulation technique. For channel estimation, pilot symbols recognized by both the transmitter and receiver are inserted alongside data symbols. In addition, they together form an OFDM waveform. The transmitted OFDM signals are transformed into parallel data streams. Then, the inverse discrete Fourier transform (IDFT) is applied to convert the OFDM signals from the frequency domain to the time domain. After that, the cyclic prefix (CP) is expanded into the OFDM symbols by replicating the previous samples and including them in the forefront of the transmitted OFDM symbols to alleviate the impact of inter-symbol interference (ISI). Lastly, the time-domain signals with CP are transformed into a serial data stream and then conveyed and propagated via wireless channels. The received signal $y(n)$ in time-domain is given by [24]:

$$y(n) = x(n) \odot h(n) + w(n), \tag{1}$$

where the operator $\odot$ denotes the circular convolution of transmitted signal $x(n)$ and the channel impulse response $h(n)$, and $w(n)$ symbolizes the additive white Gaussian noise (AWGN) with zero-mean.

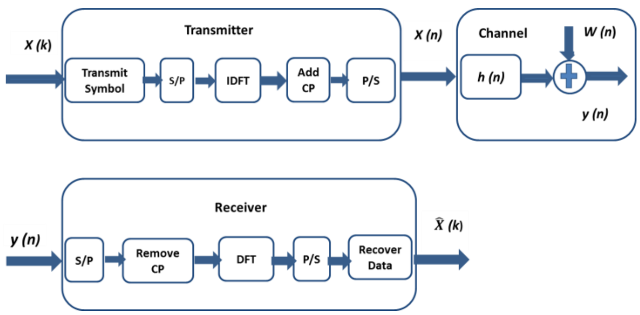

Fig. 1. OFDM system architecture.

The process is reversed on the receiving end. After parallelizing the received serial data streams, the CP is removed, and the time-domain signals are converted to the frequency-domain signals by discrete Fourier transform (DFT). Lastly, the signal is transformed back into a serial data stream for output. Hence, the received signal can be defined as:

$$Y(k) = X(k)H(k) + W(k), \tag{2}$$

where $W(k)$, $H(k)$, $X(k)$, and $Y(k)$ are the DFT of $w(n)$, $h(n)$, $x(n)$, and $y(n)$ respectively.

3. DL-based channel estimation and signal detection

This section describes in detail the architecture of the proposed DL-based CSE/SD functions. Then we briefly outline how the training phases are conducted.

3.1 Proposed DL architecture

The gated recurrent unit (GRU) network is a type of recurrent neural network (RNN) that is relatively new [25]. RNNs can be used to learn the characteristic features of time-series data and predict outcomes; however, RNNs have limited short-term memory. The LSTM network is a widely known RNN that can capture long-term relationships while also avoiding the exploding and vanishing gradient problem in long-term dependency tasks [26].

The GRU network inherits the benefits of LSTM while having a simplified architecture and fewer parameters, resulting in improved generalization ability and less computation. LSTM and GRU both have internal mechanisms known as "gates" that can control the flow of information. Two gates regulate the update of cell values in the LSTM network: the forget gate and the input gate. Due to the need for two gate architectures, the LSTM structure is relatively complicated [25, 26].

In contrast to LSTM, GRU regulates the forget parameter in addition to the update parameter values for the output through a single update gate, resulting in less computational complexity. This simplification allows the GRU to maintain LSTM functionality while reducing network training time. Unlike LSTM, GRU has only two gates: a reset gate and an update gate. Figure 2 illustrates the GRU cell structure.

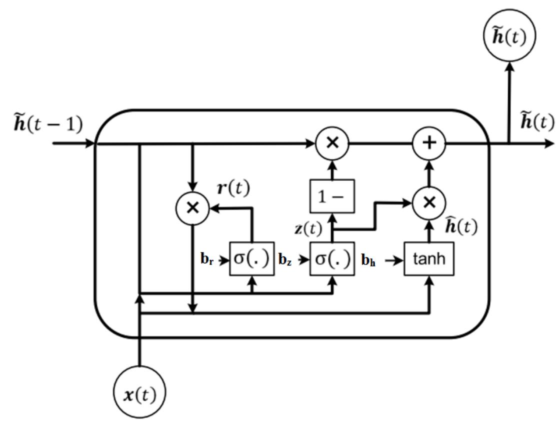

Fig. 2. Structure diagram of GRU cell.

The update gate $z(t)$ functions similarly to the forget and input gates of an LSTM. It decides what information to discard and what new information to include. Another gate that is used to decide how much past information to forget is the reset gate $r(t)$. The standard GRU architecture is specified mathematically by the following equations [27]:

$$z(t) = \sigma\left(w_z\, x(t) + U_z \tilde{h}(t-1) + b_z\right), \tag{3}$$

$$r(t) = \sigma\left(w_r\, x(t) + U_r \tilde{h}(t-1) + b_r\right), \tag{4}$$

$$\hat{h}(t) = tanh\left(w_h\, x(t) + U_h\left(\tilde{h}(t-1) \odot r(t)\right) + b_h\right), \tag{5}$$

$$\tilde{h}(t) = z(t) \odot \tilde{h}(t-1) + \left(1 - z(t)\right) \odot \hat{h}(t), \tag{6}$$

where $x(t)$ represents the current input vector and w_z, w_r, w_h (input weight matrices) and U_z , U_r , U_h (recurrent weight matrices), while b_z , b_r , b_h represent (bias vectors). $\tilde{h}(t-1)$denotes the input data at time -1 , $\tilde{h}(t)$ and $\hat{h}(t)$ represents the output and candidate states at time t. Hadamard product of vectors is represented by $\odot$. Both gates' activations are logistic sigmoid functions σ , which limit $r(t)$ and $z(t)$ to values between 0 and 1.

For CSE/SD function, we have applied DL GRU recurrent neural network architecture. The proposed DL model comprised of an input layer of the same size as the received feature vector, 256 in the current study. Then a GRU layer with 16 hidden units is used. The GRU layer's output is routed first to the fully connected layer with a size of 4 and then to the soft-max activation layer. Lastly, a classification layer classes the values from the soft-max layer to one of the mutually exclusive classes using the specified loss function (cross-entropy, MAE, or SSE). Figure 3 illustrates the layout of the proposed DL GRU model.

Figure 3 illustrates the layout of the proposed DL GRU model.

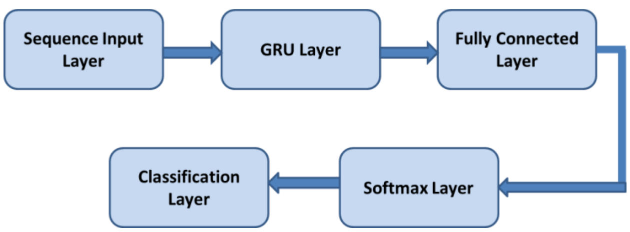

Fig. 3. The proposed DL GRU model layout with variant layers.

3.2 Training of the proposed DL model

Generally, the implementation of a DLNN model consists of two stages: the training stage and the deployment stage [23]. Before the implementation, to efficiently estimation of the channel parameters and retrieve/predict the transmitted symbols in the deployment stage, the proposed DL model must be trained on training datasets in the training phase. In the current study, training is performed offline, and deployment is done online. Figure 4 demonstrates the processes for creating training sets and performing offline DL to create a trained GRU model.

The training data set for one subcarrier is randomly generated during offline training mode. For two successive OFDM blocks, the pilot symbols are assumed to be contained within the first OFDM block, while the transmitted symbols are in the subsequent OFDM blocks. Together, these two symbols can be referred to as a frame. The channel is assumed to be constant throughout the pilot and data blocks. However, a change occurs from one frame to the next.

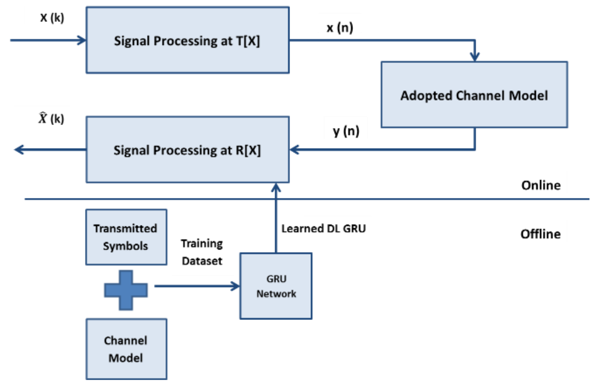

Fig. 4. Generation of training data sets for the proposed model and offline DL process.

The training datasets are created by sending OFDM frames over the adopted channel model. The necessary training dataset includes the received OFDM signals, which are influenced by existing channel characteristics and noise, as well as the originally transmitted symbols. During the online deployment stage, the proposed DL model, which was trained offline, takes the unknown received signals as input and retrieves the transmitted signals using the trained knowledge. Figure 5 shows the flowchart of training procedures for the proposed DL GRU NN.

In this study, the proposed DL GRU model is trained by the Adam optimizer in all simulation scenarios [28]. On the other hand, three various loss functions: mean absolute error (MAE), cross-entropy function for k_{th} mutually exclusive classes, and the sum of squared of the errors (SSE) are utilized to get the most effective DL GRU model. The loss functions are described as follows [29]:

$$crossentropyex = -\sum_{i=1}^{N}\sum_{j=1}^{C} X_{ij}(k)\log\left(\hat{X}_{ij}(k)\right), \tag{7}$$

$$SSE = \sum_{i=1}^{N}\sum_{j=1}^{C}\left(X_{ij}(k) - \hat{X}_{ij}(k)\right)^2, \tag{8}$$

$$MAE = \frac{\sum_{i=1}^{N}\sum_{j=1}^{C}\left|X_{ij}(k) - \hat{X}_{ij}(k)\right|}{N}, \tag{9}$$

where N represents the entire number of samples, C denotes the entire number of classes, X_{ij} denotes the i^{th} sample data sent for the j^{th} category, and $\hat{X}_{ij}$ is the proposed model's outputs for sample $i$for category j.

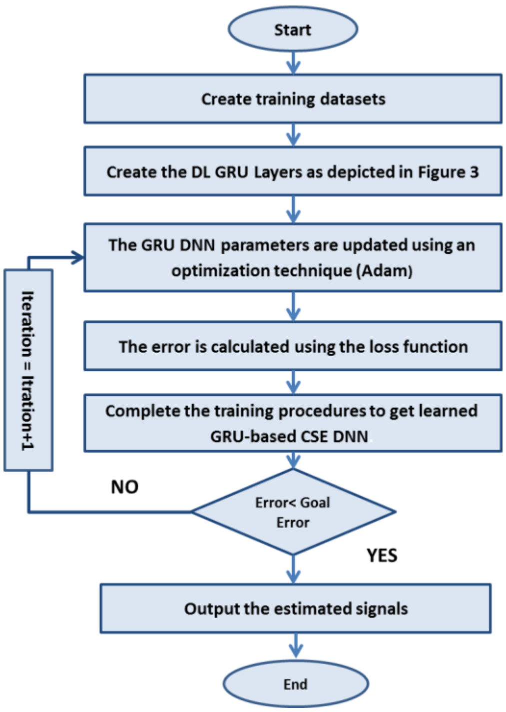

Fig. 5. The flowchart of training procedures for the proposed DL GRU NN.

4. Simulation results

In this part, extensive simulation is carried out to evaluate the performance of the proposed DL GRU-based CSE/SD. Table 1 shows the simulation parameters for the utilized OFDM system and the selected channel model, while the parameters of the proposed DL GRU framework are listed in Tab. 2. The data set for training and validation is created for a single subcarrier.

The simulated data is created utilizing a previously determined channel model with a randomly generated signal as input and transformed into an OFDM frame using the simulation settings. The simulated data is then transformed into (X_{Train}, X_{Test}) input data and (Y_{Train}, Y_{Test}) outcomes. The DL GRU model is fitted with (X_{Train}, Y_{Train}) and it is verified with the equivalent (X_{Test}, Y_{Test}). Table 3 shows the data size used in this study.

TABLE 1. Channel model and OFDM system settings

Parameter	Value
Channel Model	Rayleigh Fading
Number of Subcarriers	64
Modulation Type	Quadrature phase shift keying (QPSK)
Number of Paths	24
Carrier Frequency	2.6 GHz
Number of Pilots	64,8
Cyclic Prefix (CP) Length	16,8,0
Noise Model	Additive white Gaussian noise (AWGN)

TABLE 2. Proposed DL GRU architecture and the training parameters

Parameter	Value
Input Layer Size	256
LSTM Layer Size	16 hidden neurons
Fully Connected Layer Size	4
Number of Epochs	1000
Mini Batch Size	1000
Optimization Algorithm	Adam
Loss Function	Cross-entropy, MAE, SSE

TABLE 3. Data settings

Parameter	Value
Number of OFDM packets	10000
X_{Train}	8000 x 1 cells
Y_{Train}	8000 x 1 categorical
X_{Test}	2000 x1 cells
Y_{Test}	2000 x1 categorical

4.1 Effect of the number of pilots and cyclic prefix length on system performance

Experimental studies have been conducted to demonstrate the performance of our proposed DL GRU structure for an efficient estimate of the channel in addition to accurately retrieving transmitted symbols.

In this subsection, the performance of the proposed DL GRU framework is compared with the conventional LS and MMSE estimation techniques, in addition to the DL BiLSTM model used in [21, 22] in terms of SER versus SNR. Different cyclic prefix lengths of 16, 8, and 0 and different pilot densities of 64 and 8 will be used to evaluate the performance of the examined estimators. The proposed DL GRU model will be trained using the Adam optimizer, and the cross-entropy loss function will be applied in the final classification layer.

When 64 pilots and the length of CP of 16 are used, the proposed DL GRU-based CSE exhibits comparable performance to the examined channel estimators over a low SNR range (0–7 dB), as illustrated in Fig. 6. The proposed DL GRU framework exceeds the competitive estimation methods, starting at 8 dB.

At the length of CP of 8, Fig. 7 shows that the proposed DL GRU model outperforms the conventional estimators at all SNRs. Furthermore, starting at 15 dB, the proposed GRU model significantly outperforms the DL BiLSTM model used in [21, 22]. The traditional LS estimator, on the other hand, performs the worst.

The proposed DL GRU-based CSE outperforms the other estimators in a simulated scenario of 64 pilots without CP, as described in Fig. 8. Also, the findings indicate that the BiLSTM model is highly comparable to the MMSE estimator. The LS estimator, on the other hand, still has the worst performance.

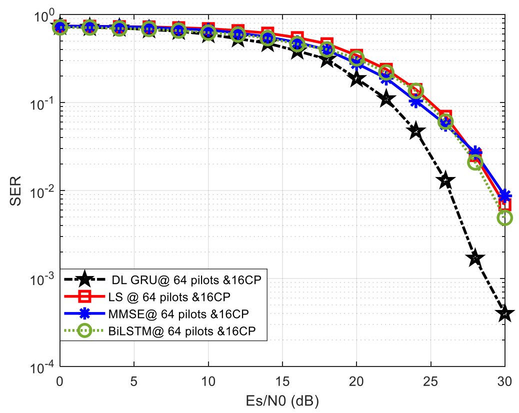

Fig. 6. SER performance curves of the proposed DL GRU model and the different estimators with 64 pilots and a CP length of 16 using the Adam optimizer and cross-entropy loss function.

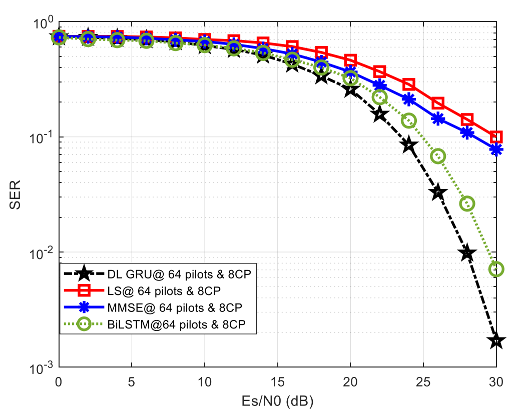

Fig. 7. SER performance curves of the proposed DL GRU model and the different estimators with 64 pilots and a CP length of 8 using the Adam optimizer and cross-entropy loss function.

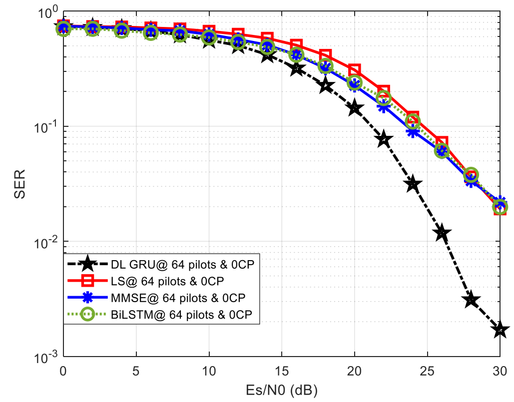

Fig. 8. SER performance curves of the proposed DL GRU model and the different estimators with 64 pilots and without CP using the Adam optimizer and cross-entropy loss function.

It is clear from Figures 6, 7, and 8 that the LS estimator always gives the worst SER performance in all cases because its estimating method relies on no prior knowledge of channel statistics. The MMSE estimator, in contrast, uses mean and covariance matrices (second-order channel statistics), resulting in better performance than its LS counterpart. In all scenarios, the SER performance of our proposed DL GRU CSE was better than that of the two traditional methods and the DL BiLSTM model used in [21, 22]. This demonstrates the effectiveness of the proposed DL GRU structure in jointly estimating the channel state and detecting transmitted symbols due to the feature of the GRU layer architecture, which allows it to remember previously processed information better than the DL BiLSTM model. Furthermore, it proves the robustness of the proposed DL GRU structure with the short/no cyclic prefix.

Figure 9 illustrates the estimating methods' behavior when the number of pilots is limited to 8, and the length of CP is 16. As this figure shows, the proposed DL GRU-based CSE significantly beats the examined estimators starting at 7 dB. Also, neither LS nor MMSE can efficiently estimate the channel information.

The proposed DL GRU estimator still provides the best performance over its counterparts when the length of CP decreases to 8, as shown in Fig. 10. On the other hand, the SER curve of the traditional MMSE and LS estimators saturated at SNR values exceeding 10 dB.

In the simulation scenario with 8 pilots and no CP, the DL GRU-based CSE still has the best SER performance compared to its peers, as shown in Fig. 11. In addition, the DL BiLSTM model performs similarly to the MMSE estimator. It is also noted that the MMSE has a better SER performance than the LS estimator, which offers the worst performance.

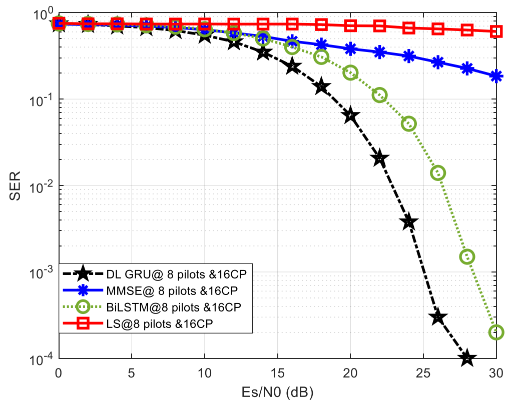

Fig. 9. SER performance curves of the proposed DL GRU model and the different estimators with 8 pilots and a CP length of 16 using the Adam optimizer and cross-entropy loss function.

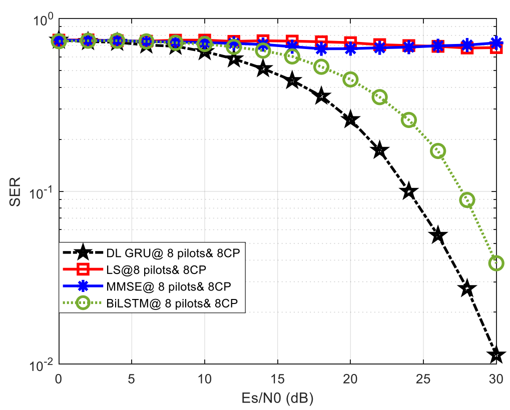

Fig. 10. SER performance curves of the proposed DL GRU model and the different estimators with 8 pilots and a CP length of 8 using the Adam optimizer and cross-entropy loss function.

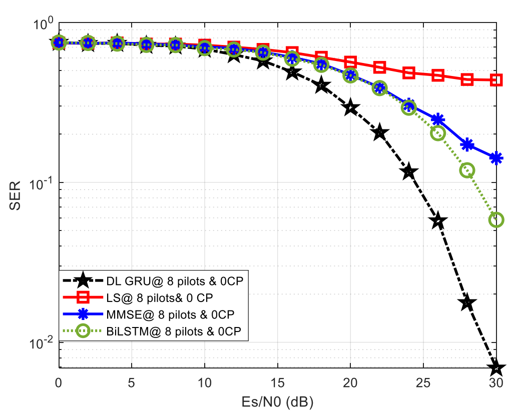

Fig. 11. SER performance curves of the proposed DL GRU model and the different estimators with 8 pilots and without CP using the Adam optimizer and cross-entropy loss function.

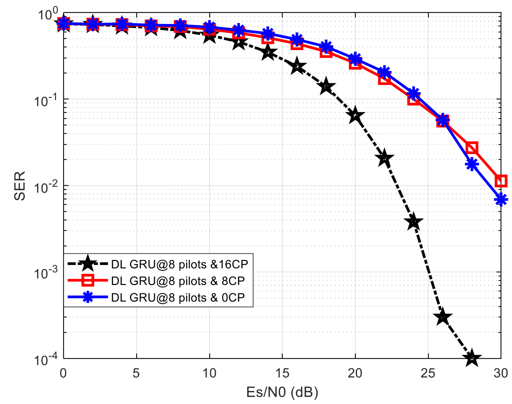

Fig. 12. SER performance of the proposed DL GRU framework at 8 pilots and CP lengths of 16, 8, and zero using the Adam optimizer and cross-entropy loss function.

The performance of the proposed DL GRU-based CSE at 8 pilots and different CP lengths of 16, 8, and 0 is summarized in Fig. 12. At low SNRs, we can see that the suggested DL GRU architecture with short/no CP has the same performance over (0–8 dB) SNRs. Furthermore, the suggested DL GRU model with CP shows fewer variances over the SNR ranges than its counterpart without CP (8–14 dB).

In summary, we can conclude from the obtained results that the proposed DL GRU-based CSE/SD is robust under the conditions of few pilots and with short/no CP. This benefit is critical for the DL CSE to implement in real time since the same performance can be reached with much fewer computations. Furthermore, for OFDM wireless communication systems, the proposed DL GRU structure with minimal spectrum resources used for channel state estimation is recommended to significantly improve their spectrum efficiency, energy efficiency, and transmission data rates.

4.2 Impact of various loss functions on system performance

The loss function is essential to creating and improving the performance of the DL algorithms. The lower the value of the loss function, the better performance is achieved. Generally, the DL model is trained using an optimizer that employs a loss function to compute an error between the expected output of the model and the predicted outcome. There are numerous loss functions, and selecting the proper one for a given problem might be difficult. During the training stage, learning algorithms (optimizers) attempt to minimize the current loss function to the specified error target by iteratively optimizing the DLNN weights and biases at each training epoch.

The current subsection investigates the performance of the proposed DL GRU structure with various loss functions at 8 pilots and different CP lengths of 16, 8, and 0. Cross-entropy, which was investigated in the previous subsection, MAE, and

SSE are the three loss functions utilized in the classification layer of the proposed DL GRU model. The Adam optimization algorithm is employed to train the proposed DL GRU network, and various loss functions are applied in the final layer to obtain the most effective and robust version of the proposed DL GRU-based CSE/SD.

Figure 13 indicates that at 8 pilots, CP lengths of 16, and with Adam optimization, the proposed DL GRU structure with a cross-entropy-based classification layer model achieves the same performance as the SSE- and MAE-based classification layer models over a low SNR range (0-10 dB) and (0-7 dB), respectively. Above these SNR levels, the proposed DL GRU structure with a cross-entropy-based classification layer model beats both the SSE- and MAE-based classification layer models. In contrast, the MAE model shows the lowest performance.

The proposed DL GRU structure with the SSE-based classification layer model provides superior performance when the length of CP decreases to 8, as shown in Fig. 14. Furthermore, the cross-entropy model exceeds the MAE model.

In the simulation scenario of 8 pilots and the absence of the CP, the proposed DL GRU with (cross-entropy, SSE, or MAE)-based classification layer models show comparable performance over the SNR range (0–20 dB), as illustrated in Fig. 15. Beyond these SNR ranges, the cross-entropy model outperforms the SSE and MAE models.

Figure 16 confirms the robustness of the proposed DL GRU model with a limited number of pilots against the very short/without CP used. Moreover, they highlight the significance of analyzing different loss function-based classification layers during the deep learning process to attain the most effective model of any suggested DL GRU structure.

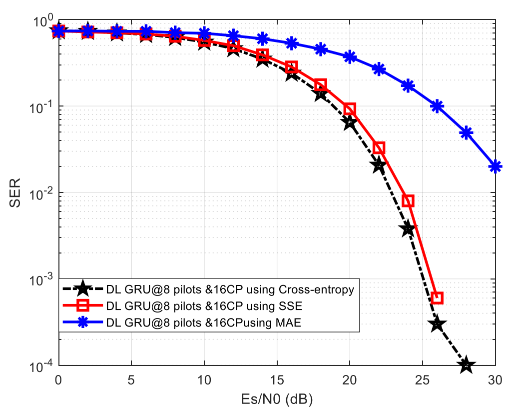

Fig. 13. SER performance of the proposed DL GRU model at 8 pilots and a CP length of 16 by using the Adam optimizer and various loss functions.

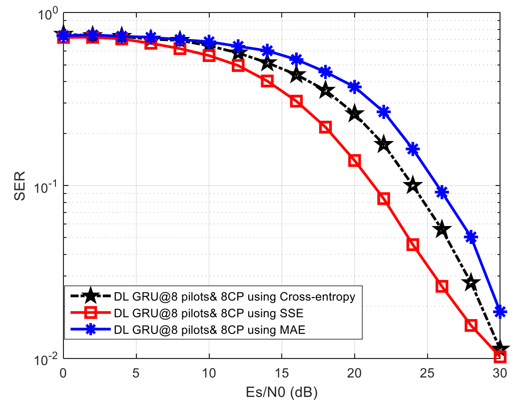

Fig. 14. SER performance of the proposed DL GRU model at 8 pilots and a CP length of 8 by using the Adam optimizer and various loss functions.

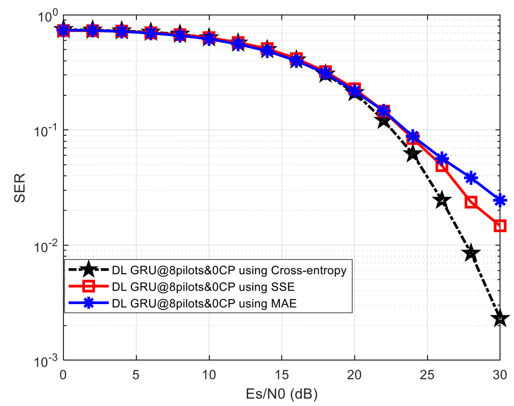

Fig. 15. SER performance of the proposed DL GRU model at 8 pilots and without CP using the Adam optimizer and various loss functions.

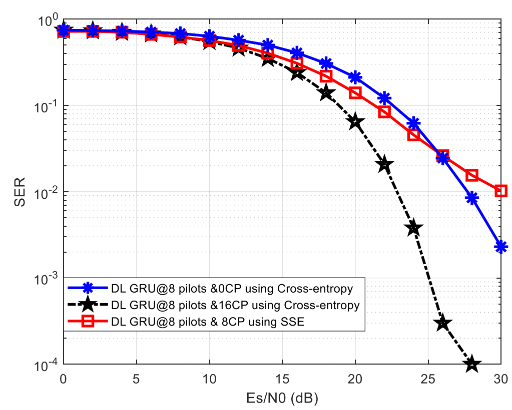

Fig. 16. SER performance comparison of the best GRU-based DL model with a limited number of pilots (8 pilots) using different lengths of CP and various loss functions.

5. Conclusion

In this study, a DL architecture based on GRU recurrent neural networks has been proposed for the channel state estimation (CSE) and signal detection (SD) applications. The proposed DLNN performs in an end-to-end manner and combines CSE and SD functions in OFDM wireless communication systems. The proposed DL GRU framework has been trained in an offline manner using the received OFDM signals that have been exposed to various channel defects before being executed online to extract/retrieve the transmitted data symbols. Several investigations have been carried out to assess the performance of the proposed DL GRU structure and show its efficiency for CSE/SD functions compared to the classical LS and MMSE estimation methods, in addition to the DL-based BiLSTEM model. The simulation results indicate that the proposed DL GRU model beats the traditional LS and MMSE estimators as well as the DL BiLSTM model in terms of SER. They also show that the proposed DL GRU RNN framework has a lot of potential for CSE and SD while reducing computational requirements. Furthermore, the proposed DL GRU framework can adapt to a reduction in pilots' number and CP length that traditional approaches cannot. Because the proposed DL-based technique, which employs a GRU RNN, is data-driven, has fewer computational requirements, and is independent of channel characteristics; it could be applied to any channel situation. In short, the suggested DL GRU framework offers a potential alternative for CSE/SD in OFDM wireless communication systems. On the other hand, the simulation results prove the significance of examining different loss functions to acquire the most effective model of the proposed DL GRU framework. In future work, the proposed DL GRU-based CSE/SD will be used in more complicated system models, such as MIMO environments.

References

[1] T. Hwang, C. Yang, G. Wu, S. Li and G. Li, “OFDM and its wireless applications: A survey”, *IEEE Trans*. Vehicular Technol., vol. 58, no. 4, pp. 1673-1694, 2009.

[2] Tao Cui and C. Tellambura, “Joint data detection and channel estimation for OFDM systems”, *IEEE Transactions on Communications*, vol. 54, no. 4, pp. 670-679, 2006.

[3] Y. Li, J. Winters and N. Sollenberger, “MIMO-OFDM for wireless communication: Signal detection with enhanced channel estimation”, *IEEE Trans. Commun.*, pp. 1471-1477, 2002.

J. K. Moon and S. I. Choi, “Performance of channel estimation methods for OFDM systems in a multipath fading channels”, IEEE Trans. Consum. Electron, vol. 46, no. 1, pp. 161-170, 2000.

[5] M. K. Ozdemir and H. Arslan, “Channel estimation for wireless OFDM systems”, *Commun. Surveys Tuts.*, vol. 9, no. 2, pp. 18-48, Second Quarter 2007.

[6] S. Coleri, M. Ergen, A. Puri and A. Bahai, “Channel estimation techniques based on pilot arrangement in OFDM systems”, *IEEE Trans. Broadcast.*, vol. 48, no. 3, pp. 223-229, 2002.

[7] Tian-Ming Ma, Yu-Song Shi, and Ying-Guan Wang, “A Low Complexity MMSE for OFDM Systems over Frequency-Selective Fading Channels”, *IEEE Communications Letters*, vol. 16, no. 3, 2012.

[8] S. M. Aldossari and K.-C. Chen, “Machine learning for wireless communication channel modeling: an overview”, *Wireless Personal Communications*, vol. 106, no. 1, pp. 46–70, 2019.

[9] W. Liu, Z. Wang, X. Liu, N. Zeng, Y. Liu, and F. E. Alsaadi, “A survey of deep neural network architectures and their applications”, *Neurocomputing*, vol. 234, pp. 11–26, 2017.

[10] Zhang, Chaoyun, Paul Patras, and Hamed Haddadi, “Deep learning in mobile and wireless networking: A survey”, *IEEE Communications surveys & tutorials*, vol.21, no.3, pp. 2224-2287, 2019.

[11] L. Dai et al., “Deep learning for wireless communications: An emerging interdisciplinary paradigm”, *IEEE Wireless Commun.*, vol. 27, no. 4, pp. 133–139, 2020.

[12] H. Huang et al., “Deep Learning for Physical-Layer 5G Wireless Techniques: Opportunities, Challenges and Solutions”, *IEEE Wireless Commun.*, vol. 27, no. 1, pp.214–222, 2020.

[13] Le. H. A., Van. Chien. T., Nguyen. T. H., Choo. H. and Nguyen. V. D, “Machine Learning-Based 5G-and-Beyond Channel Estimation for MIMO-OFDM Communication Systems”, *Sensors*, vol. 21, no. 14, 2021.

[14] A. K. Gizzini, M. Chafii, A. Nimr and G. Fettweis, “Deep Learning Based Channel Estimation Schemes for IEEE 802.11p Standard”, *IEEE Access*, vol. 8, pp.113751–113765, 2020.

[15] R. Jiang, X. Wang, S. Cao, J. Zhao and X. Li, “Deep Neural Networks for Channel Estimation in Underwater Acoustic OFDM Systems”, *IEEE Access*, vol. 7, pp.23579–23594, 2019.

[16] M. H. Essai Ali, “Deep learning-based pilot-assisted channel state estimator for OFDM systems”, *IET Communications*, vol. 15, no. 2, pp. 257-264, 2021.

L. Li, H. Chen, H.-H. Chang, and L. Liu, “Deep residual learning meets OFDM channel estimation”, *IEEE Wireless Commun. Lett.*, vol. 9, no. 5, pp. 615-618, 2020.

[18] Z. Zhao, M. C. Vuran, F. Guo and S. D. Scott, "Deep-Waveform: A Learned OFDM Receiver Based on Deep Complex-Valued Convolutional Networks", *IEEE Journal on Selected Areas in Communications*, vol. 39, no. 8, pp. 2407-2420, 2021.

[19] Liao, Yong, et al., "ChanEstNet: A deep learning based channel estimation for high-speed scenarios", in *Proc. IEEE Int. Commun. Conf. (ICC)*, Shanghai, China, May 2019, pp. 1–6.

[20] G. Pan, Z. Liu, W. Wang and M. Li, "A Signal Detection Scheme Based on Deep Learning in OFDM Systems", *IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pp. 1–6, 2021.

[21] S. Wang, R. Yao, T. A. Tsiftsis, N. I. Miridakis and N. Qi, "Signal Detection in Uplink Time-Varying OFDM Systems Using RNN With Bidirectional LSTM", *IEEE Wireless Communications Letters*, vol. 9, no. 11, pp. 1947-1951, Nov. 2020.

[22] A. K. Nair and V. Menon, "Joint Channel Estimation and Symbol Detection in MIMO-OFDM Systems: A Deep Learning Approach using Bi-LSTM," *14th International Conference on Communication Systems & NETworkS (COMSNETS)*, pp. 406-411, 2022.

Huang, S. C., and Le, T. H., *Principles and Labs for Deep Learning*, Academic Press: Cambridge, MA, USA, 2021.

[24] H. Ye, G. Y. Li, and B. Juang, "Power of Deep Learning for Channel Estimation and Signal Detection in OFDM Systems", *IEEE Wireless Communications Letters*, vol. 7, no. 1, pp. 114-117, 2018.

[25] Dey, Rahul, and Fathi M. Salem, "Gate-variants of gated recurrent unit (GRU) neural networks", *IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*, pp. 1597-1600, 2017.

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory", *Neural Comput.*, vol. 9, no. 8, pp. 1735-1780, 1997.

[27] M. Ravanelli et al., "Light gated recurrent units for speech recognition", *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 2, no. 2, pp. 92-102, 2018.

[28] Jais, Imran Khan Mohd, Amelia Ritahani Ismail, and Syed Qamrun Nisa, "Adam optimization algorithm for wide and deep neural network", *Knowledge Engineering and Data Science*, vol. 2, no. 1, pp. 41-46, 2019.

[29] M. H. Essai and I. B. Taha, "Channel state information estimation for 5G wireless communication systems: recurrent neural networks approach", *PeerJ Computer Science* 7, 2021, e682.

Received 9 March 2023