Deep learning-based pilot-assisted channel state estimator for OFDM systems

Mohamed Hassan Essai Ali

Electrical Engineering Department, Al-Azhar University, Qena, Egypt

Correspondence
Mohamed Hassan Essai Ali, Electrical Engineering Department, Al-Azhar University, Qena, Qena, Egypt.
Email: mhessai@azhar.edu.eg

Abstract
This study proposes an online deep learning-based channel state estimator for OFDM wireless communication systems by employing the deep learning long short-term memory (LSTM) neural networks. The proposed algorithm is a pilot-assisted estimator type. The proposed estimator is initially offline trained using simulated data sets, and then it follows the channel statistics in an online deployment, where finally the transmitted data can be recovered. A comparative investigation is performed using three different optimisation algorithms for deep learning to evaluate the performance of the proposed estimator at each. The proposed estimator provides a superior performance in comparison to least square (LS) and minimum mean square error (MMSE) estimators when limited pilots are used, thanks to the outstanding learning and generalisation capabilities of deep learning LSTM neural networks. Also, it does not require any prior knowledge of channel statistics. So, the proposed estimator is promising for channel state estimation in OFDM communication systems.

1 | INTRODUCTION
Orthogonal frequency-division multiplexing (OFDM) is a common modulation method that has been approved in wireless wideband systems to mitigate frequency-selective fading in wireless channels. The quality of channel response estimation is crucial to the performance of the OFDM wireless communication systems.

Channel estimators can be categorised as pilot-assisted, blind, and decision-directed estimators. Pilot-assisted channel state estimation is the most common approach where a transmitter sends a pilot, which acts as the reference signal that is used by a sending and a receiving ends. Thanks to its very low computational complexity, pilot-assisted channel estimators can be applied in any wireless communication system. However, its main drawback is a decrease in the transmission rate, since pilot signals are inserted. Thus, one of the most design challenges of pilot-assisted channel estimator is to minimise the number of pilots and accurately estimate the state of the channel [1].

Blind channel estimator does not require pilots where it uses inherent information in the received symbols; unfortunately, it suffers from high computational complexity and latency in contrast to pilot-assisted channel estimators. Thus, blind channel estimation is infrequently adopted in practical wireless communication systems.

Decision-directed estimators use pilot and detected data symbols and hence update the channel estimations. Thus, they achieve superior performance compared to the pilot-assisted channel estimator. One of the substantial design concerns for channel estimator in OFDM systems is channel estimator design with both low complexity and high accuracy at low pilots’ number.

In this study, focus is on the pilot-assisted channel estimation in OFDM systems using deep learning (DL) long short-term memory (LSTM) neural networks, referred as DLLSTM-based channel state estimator (CSE) in the rest of the study. To the best of my knowledge, it is the first time that the LSTM neural network is to be used as a CSE. In [2–4], the conventional feed forward neural networks (FFNN) and 1D-convolutional neural network (1D-CNN) have been employed respectively to build CS estimators. Also, the proposed estimator outperforms the conventional LS and minimum mean square error (MMSE) estimators at the use of a limited number of pilots.
There are three reasons for the use of deep learning neural networks (DLNNs) in various fields [5, 6]. First, DLNN-based CSEs are data-based and therefore more resistant to imperfections in real systems. Second, DLNN-based CSEs have low computational complexity, which includes only a few levels of simple matrices and vectors operations. Third, with the rapid enhancement of the parallel processing capabilities of the specialised chips such as the graphic processing units (GPU) [7], the implementation of DLNNs can be easily parallelised on parallel architectures and easily implemented with low data types accuracy, which makes DLNN-based approaches much more efficient. Substantiated by these advantages, DLNNs were represented at the physical level and achieved outstanding performance in various fields [8].

Due to the long training time of the proposed DLLSTM-based CSE and the large numbers of the weights and other parameters that must be updated and adjusted during the learning process, it will be trained offline. Then the trained DLLSTM-based CSE is used in online employment to recover the transmitted data.

The proposed DLLSTM-based CSE can successfully learn and analyse the characteristics of the wireless channels that may suffer from distortions and interferences. The performance of the proposed DLLSTM-based CSE will be compared with the performance of the most commonly used least square (LS) and MMSE CSEs. Also, the proposed estimator will be trained using three different optimisation algorithms, using the collected simulation datasets, in order to get the most efficient estimator model with lower pilots.

The obtained results show that the DLLSTM-based CSE achieves performance comparable to the traditional MMSE estimator at high enough pilots, and it outperforms both LS and MMSE estimators at limited pilots and channel interference. Also, the results indicate that the DLLSTM-based CSE can be potentially applied in OFDM wireless communications systems with a limited number of pilots in order to enhance the system transmission rate.

The rest of this study is organised as follows. Deep neural networks-based CSE is introduced in Section 2. OFDM communication system and the proposed DLLSTM-based CSE is presented in Section 3. Simulation results are presented in Section 4. Conclusions are given in Section 5.

2  |  DLNN-BASED CSE

In this section, LSTM neural networks is presented for channel estimation. The proposed DLLSTM-based CSE is trained offline using the simulated data, regardless what the OFDM and the wireless channel are.

The LSTM network is a recurrent neural network that has the ability to learn the long-term relationships between the time steps of sequence data [9]. Many LSTM-based approaches have been developed to solve problems such as handwriting recognition [10], speech recognition [11], online translation such as Google neural machine translation [12] and Facebook translation system [13].

In the realm of wireless communications, intelligent mechanisms, especially artificial neural networks (feed forward, or recurrent neural networks), have already been adopted to CSI-based localisation [14], channels decoding [15], data traffic, user location, channel load, and service requests [16], channel estimation and detection [2, 3]. In [2], the authors proposed FFDDLNN-based joint channel estimation and symbol detection approach for (OFDM) systems with frequency selective channels. The proposed algorithm is shown to outperform the conventional MMSE estimator when imperfections of the communication systems are considered. In [3], the authors proposed an online feedforward DL-based estimator for doubly selective channels. The proposed algorithm is shown to outperform the conventional LMMSE estimator in all examination scenarios. In [4], a 1D-CNN DL model was proposed to estimate the provided channel and recover the equalised data. Also, the performance of 1D-CNN was compared with the LS, MMSE, and FFNN estimators in terms of BER and MSE at different modulation techniques. 1D-CNN is shown to outperform both LS, MMSE, and FFNN estimators.

In the current study, the proposed DLLSTM-based CSE using LSTM neural network is being constructed. The proposed CSE will be trained using different three DLNN optimisation algorithms. These are adaptive moment estimation (Adam), root mean square propagation (RMSProp), and stochastic gradient descent with momentum (SGDm), in order to get the most reliable and robust estimator under the conditions of limited pilots’ numbers.

To construct DL LSTM neural network for the task of channel state estimation, an array of the following five layers has been created: A sequence input layer, an LSTM layer, a fully connected layer, a softmax layer, and a classification output layer; input size as 256 (the number of features of the input data), LSTM layer to have 16 hidden units and to output the last element of the sequence; finally, four classes by including a fully connected layer of size 4, followed by a softmax layer and a classification layer. Deeper LSTM networks can be built by inserting extra LSTM layers. Figure 1(a) depicts the structure of the proposed CS estimator. The five blocks in Figure 1(a) can be grouped together into a single compacted block as in Figure 1(b).
3 | OFDM COMMUNICATION SYSTEM AND THE PROPOSED DLLSTM-BASED-CSE

In the following subsections, the conventional OFDM communication system and offline DL of DLLSTM-based CSE are introduced briefly.

3.1 | OFDM system model

The architecture of the conventional OFDM communication system is depicted in Figure 2.

On the transmitter end, the transmitted data (symbols) with pilots are initially converted to a parallelized data stream using a serial-to-parallel (S/P) converter, then the signal is converted from the frequency domain to the time domain using the inverse discrete Fourier transform (IDFT). In order to alleviate the effect of the inter-symbol interference (ISI), the cyclic prefix (CP) must be inserted. The maximum delay spread of the channel should be shorter than the CP length.

As in [2], the multipath channel of a sample space described by complex random variables \( \{ b(n) \}_{n=0}^{N-1} \) will be considered, thus the received signal can be represented by

\[
y(n) = x(n) \oplus b(n) + w(n)
\]

where \( \oplus \) is the circular convolution, \( x(n) \) is the input signal, \( y(n) \) is the received (observed) signal, \( w(n) \) and is the additive white Gaussian noise (AWGN).

The received signal in the frequency domain can be expressed as

\[
Y(k) = X(k)H(k) + W(k)
\]

where \( Y(k) \), \( X(k) \), \( H(k) \), and \( W(k) \) are the discrete Fourier transformations of \( y(n) \), \( x(n) \), \( b(n) \), and \( w(n) \), respectively. These DF transformations were obtained after removing the cyclic prefix.

OFDM frame consists of pilot symbols in the first OFDM block, followed by the transmitted data in the next OFDM blocks. The channel can be considered stationary over a specific frame but change between the different frames. The proposed DLLSTM-based CSE accepts the received data at its input and retrieves the transmitted data at its output.

3.2 | Offline DL of DLLSTM-based CSE

Recently, researchers have developed many channel models for CSI that well characterise the channel statistics of real channels. Using these channel models, training data can be obtained by modelling. The current study adopts 5G channel model. This model is described in the 3GPP document TR38.901 [17, 7] to mimic the behaviour of the channel model that have many imperfections and unknown effects that degrade the performance of the CS estimators. Other channel models can be used such as the narrowband Rayleigh fading channel and doubly selective fading channels [3].

In offline training, the OFDM frame (pilot, and transmitted symbols) is formed from the randomly generated data sequence. The current random (CS) is modelled based on the adopted channel model. The received OFDM signal is retrieved based on OFDM frames subjected to current channel distortion and noise. Both the transmitted and the received signals constitute the offline training data sets.

The proposed DLLSTM-based CSE is trained to minimise the loss function, that is, the difference between the response of the proposed estimator and the original transmitted data by iteratively updating the randomly generated weights and biases. The loss function can be expressed by several methods such as MSE and mean absolute error (MAE). In this study, the cross-entropy function for \( k \) mutually exclusive classes (crossentropyex) loss function has been chosen.

\[
crossentropyex = -\sum_{i=1}^{N} \sum_{j=1}^{C} X_{ij}(k) \log (\hat{X}_{ij}(k)) \quad (3)
\]

where \( N \) is the total number of samples, \( C \) is the total number of classes, \( X_{ij} \) is the \( i \)th transmitted data sample belongs to the \( j \)th class, \( \hat{X}_{ij} \) is the DLLSTM-based CSE response for sample \( i \) for class \( j \).

In this study, the most commonly used optimisation algorithms are adopted in the context of DL and the proposed estimator illustrated in Section 4.2. More information about the widely used optimisation algorithms for DL can be found in [18].

Figure 3 depicts pictorially the training data formation and offline DL process for obtaining the learned DLLSTM-based CSE.

4 | SIMULATION RESULTS

Several experiments have been conducted to demonstrate the performance of the proposed CS estimator for the OFDM wireless communication systems. The proposed estimator was trained based on the collected simulation data sets and compared with the conventional LS and MMSE CSEs in terms of
symbol error rates (SERs) under various signal-to-noise ratios (SNRs).

The training dataset is collected for one subcarrier. The transmitter sends OFDM packets to the receiver where each OFDM packet contains one OFDM pilot symbol and one OFDM data symbol. In the pilot sequence, the data symbols may be interleaved.

Table 1, summarises the DLLSTM neural network architecture parameters and training options. Table 2 lists the OFDM system and channel parameters.

Also, in the current simulations, different optimisation algorithms will be used to train the proposed estimator, to investigate its performance at these learning approaches. These optimisation algorithms are the SGDm, RMSProp, and Adam [19].

### Table 1: Summary of the LSTM network architecture parameters and training options

| Parameter                  | Value                  |
|----------------------------|------------------------|
| Size of input              | 256                    |
| Size of LSTM layer         | 16 hidden neurons       |
| Size of the full connected layer | 4                     |
| Loss function              | crossentropy           |
| Size of mini batch         | 1000                   |
| No. of Epochs              | 1000                   |
| Optimisation algorithms    | Adam, RMSProp, and SGDm|

### Table 2: Summary of OFDM system and channel parameters

| Parameter                  | Value                  |
|----------------------------|------------------------|
| Modulation mode            | QPSK                   |
| Carrier frequency          | 2.6 GHz                |
| No. of paths               | 24                     |
| Cyclic prefix length       | 16                     |
| No. of subcarriers         | 64                     |
| No. of pilots              | 4, 8, and 64           |

### Figure 3
Training data formation and offline DL process of the proposed estimator

### Figure 4
Symbol error rate (SER) curves of the proposed and traditional estimators using 64 pilots

### Figure 5
SER curves of the proposed and traditional estimators using 8 pilots

### 4.1 Number of pilots effects

In this section, the performance of the proposed estimator with the LS and MMSE estimators will be compared. The performance of the three estimators will be evaluated at various pilots of 4, 8, and 64. In the current simulation, Adam learning algorithm will be used.

At a sufficiently high enough number of pilots, the proposed estimator has much better performance than the LS estimator and comparable performance to the MMSE estimator at SNR from 0 to 11 dB. Also, the MMSE estimator outperforms the LS estimator at all SNR examination range as shown in Figure 4.

From Figures 5 and 6, as the number of pilots gets lower 8 and 4, the proposed estimator outperforms both LS and MMSE estimators. When only 8 pilots are used, the LS and MMSE estimators have worse performance compared with the proposed...
To imperfections in real-world communication systems and achieves the best performance at 4-pilots.

4.2 Effect of optimisation algorithms on the performance of the proposed estimator

Optimisation algorithms play a vital role in improving DL processes. Training of deep neural networks can be described as an optimisation problem that seeks to find a global optimum through a reliable training trajectory and fast convergence using gradient descent algorithms [19]. The goal of a DL process is to find a model that will produce better and faster results through weights and biases adjusting to minimise the loss function (gradient descent).

Choosing the optimal optimisation approach for a specific scientific problem acts as a serious challenge. Choosing an inappropriate optimisation approach may lead the network to reside in the local minima during training, and this does not achieve any advances in the learning process. Hence, the investigation is necessary to analyse the performance of different optimisers depending on the model and dataset employed for obtaining DLLMST-based CSE with the best performance.

This section introduces an experimental comparison of the performance of three commonly used first-order SGD optimisation algorithms on the proposed DLLMST-based CSE using the collected datasets. It will also investigate how well each optimiser can deal with the problem of channel state estimation, hence finding the most efficient CSE.

The three optimisation algorithms used are Adam, which was examined in Section 4.1, RMSProp, and SGDm. Here, the learning process efficiency of both RMSProp and SGDm for obtaining a more reliable DNN-based CSE will be investigated.

It is obvious from Figures 8 and 9 that RMSProp models outperform both Adam and SGDm models at pilots of 64 and 8. At limited pilots of 4, the Adam model outperforms its
peers in terms of SER as shown in Figure 10. We can note that the performance of the same optimiser differs depending on the number of pilots. Finally, we can note that the SGDm models have the worst performance at different pilot numbers.

Figure 11 emphasizes the robustness of the proposed estimator against the used limited pilots, also it demonstrates the importance of investigating various optimisation algorithms in the DL process of the proposed estimator. It is clear that the performance of the proposed algorithm which uses 4 pilots coincides with its performance at 64 pilots, therefore using the proposed estimator with 4 pilots will be preferred for communication systems than that uses 64 pilots. This in turn will increase the transmission rate of the OFDM wireless communication systems.

During the training process of the DLLNs, it is beneficial to monitor the training processes. We can learn how the training is progressing by plotting loss metrics during training. Figures 12–14 illustrate that SGDm optimisation algorithm achieves the higher loss (the worst performance) in comparison with Adam and RMSProp optimisation algorithms, which can be verified from Figures 8–10, where the trained DLLMST-based CS estimators using SGDm algorithm have the highest SER values.

Also, the loss of both Adam and RMSProp optimisation algorithms emphasise the obtained results in Figures 8, 9, and 10 (but the reader has to zoom in the Figures 12, 13, and 14). Loss function figures can be accessed in Matlab figure format from [21].
In this study, an online DL-based CSE for OFDM systems has been proposed. The use of DLLSTM neural networks has been adopted. The proposed estimator is initially trained offline, then exploited online in the communication system to follow the channel statistics, and finally the CS can be estimated and the transmitted data can be recovered. The performance of the proposed estimator is examined at different pilots 64, 8, and 4. Also, a comparative study is performed using three different optimisation algorithms for DL to evaluate the performance of the proposed estimator at each. The proposed estimator outperforms both LS and MMSE estimators in terms of efficiency (as the SNR increases, the related SER decreases), and robustness (proposed estimators with limited pilots achieve the same performance as with high pilots’ number). Thanks to the outstanding learning and generalisation capabilities of DL LSTM neural networks, which do not require any prior knowledge of channel statistics, the proposed estimator is promising for channel state estimation in OFDM communication systems, especially those with limited pilots.

For future studies, the following is suggested.

1. Studying the performance of the proposed estimator using other different optimisation algorithms such as Adadelta, Adagrad, AMSgrad, AdaMax, and Nadam.
2. Studying the computational complexity of the proposed estimator.
3. Developing more robust loss functions using robust statistics estimators such as Huber, Cauchy, and using them to enhance the performance of the proposed estimator instead of ‘crossentropyex’ loss function under the conditions of real systems channels.

REFERENCES
1. Kim, H.: Wireless Communications Systems Design (1st edition). John Wiley & Sons Ltd., Chichester, UK (2015)
2. Ye, H., et al.: Power of deep learning for channel estimation and signal detection in OFDM systems. IEEE Wireless Commun. Lett. 7(1), 114–117 (2018)
3. Yang, Y., et al.: Deep learning-based channel estimation for doubly selective fading channels. IEEE Access 7, 36579–36589 (2019). https://doi.org/10.1109/ACCESS.2019.2901066
4. Ponnaluru, S., Penke, S.: Deep learning for estimating the channel in orthogonal frequency division multiplexing systems. J. Ambient Intell. Hum. Comput. (2020). https://doi.org/10.1007/s12652-020-02010-1
5. O’Sheaand, T., Hoydis, J.: An introduction to deep learning for the physical layer. IEEE Trans. Cognit. Commun. Networking 3(4), 563–575 (2017)
6. Jacobs, I.S., Bean, C.P.: Fine particles, thin films and exchange anisotropy. In: Rado, G.T., Suhl, H. (eds.) Magnetism, vol. 3, pp. 271–350. Academic Press, New York (1963)
7. Varrall, G.: 5G and Satellite Spectrum, Standards, and Scale. Artech House, Boston (2018)
8. Wang, T., et al.: Deep learning for wireless physical layer: Opportunities and challenges. China Commun. 14(1), 92–111 (2017)
9. Sepp, H., Jürgen, S.: Long short-term memory. Neural Comput. 1735–1780 (1997)
10. Graves, A., et al.: A novel connectionist system for improved unconstrained handwriting recognition. IEEE Trans. Pattern Anal. Mach. Intell. 31(5), 855–868 (2009)
11. Sak, H., et al.: Long short-term memory recurrent neural network architectures for large scale acoustic modeling. Archived from the original on 24 Apr 2018. https://static.googleusercontent.com/media/research.google.com/en//pbarchive/3995.pdf
12. Wu, Y., et al.: Google’s neural machine translation system: bridging the gap between human and machine translation. (2016). arXiv:1609.08144
13. Ong, T.: Facebook’s translations are now powered completely by AI. (2017). Retrieved (2019) from www.walthingsdistributed.com
14. Li, Y.G., et al.: Robust channel estimation for OFDM systems with rapid dispersive fading channels. IEEE Trans. Commun. 46(7), 902–915 (1998)
15. Wang, X., et al.: CSI-based fingerprinting for indoor localization: A deep learning approach. IEEE Trans. Veh. Technol. 66(1), 763–776 (2017)
16. Li, R., et al.: Intelligent 5G: When cellular networks meet artificial intelligence. IEEE Wireless Commun. 24(5), 175–183 (2017)
17. 3GPP TR 38.901 V16.1.0 (2019-12). Study on channel model for frequencies from 0.5 to 100 GHz (release 16), Tech. Rep. 3GPP TR V16.1.0. 3rd Generation Partnership Project (3GPP), Technical Specification Group Radio Access Network, Valbonne, France (2019)
18. Soydaner, D.: A comparison of optimization algorithms for deep learning. Int. J. Pattern Recognit. Artif. Intell. 205–213 (2020). https://doi.org/10.1142/S0218001420520138

19. Dogo, E.M., et al.: A comparative analysis of gradient descent-based optimization algorithms on convolutional neural networks. In: International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, pp. 92–99 (2018). https://doi.org/10.1109/CTEMS.2018.8769211

20. Chen, Y.H., et al.: Eyeriss: An energy-efficient reconfigurable accelerator for deep convolutional neural networks. IEEE J. Solid-State Circuits 52(1), 127–138 (2017)

21. https://azu-my.sharepoint.com/:f:/g/personal/mhessai_azhar_edu_eg/EiEeSWHBK_FKcVgOREJziVQBw5zLXb95yEAGx3f79QTUG?e=COtbgY

How to cite this article: Essai Ali MH. Deep learning-based pilot-assisted channel state estimator for OFDM systems. IET Commun 2021;15:257–264. https://doi.org/10.1049/cmu2.12051