Doppler Shift Compensation Using an LSTM-based Deep Neural Network in Underwater Acoustic Communication Systems

Sabna Hassan*, Peng Chen*, Yue Rong*, Kit Yan Chan*

*School of Electrical Engineering, Computing and Mathematical Sciences, Curtin University, Australia

OCEANS 2023 - Limerick | 979-8-3503-3226-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/OCEANSLimerick52467.2023.10244541

***Abstract*—In underwater acoustic (UA) communication, Doppler effect is much larger than that of the terrestrial radio communication as propagation speed of the acoustic wave in water is far slower than that of the electromagnetic wave in the air. Frequency offset and Doppler effect severely degrade the performance of orthogonal frequency-division multiplexing (OFDM) systems. This paper presents a Doppler shift mitigation approach in UA OFDM communication systems incorporated with a deep neural network (DNN) based receiver. The regression-based DNN is incorporated with a long short-term memory (LSTM) layer, which has an improved feature extraction capability compared with the commonly used neural network methods. The DNN is trained by simulation data with Doppler shifts and is used to predict the transmitted bits. The results show that the proposed LSTM-based DNN receiver achieves better performance than the traditional receiver, which is implemented with the least-squares (LS) channel estimator, and the commonly used neural network methods.**

***Index Terms*—Underwater acoustic communication, orthogonal frequency-division multiplexing, long short-term memory, deep neural network, convolutional neural network, Doppler shift, least-squares, regression.**

I. Introduction

The underwater acoustic (UA) channel is a time-varying, multipath complex communication channel. Reliable communication through the UA channel is challenging, due to the time-varying instantaneous frequencies caused by the propagation of sound signals, namely the Doppler effect. Doppler effect on received signals is caused by relative motion between transceiver nodes in mobile UA communication. It compresses or expands the signal waveform and shifts the frequency of the signal by generating offset [1]. The Doppler effects can cause a high error rate, and even a complete failure of the communication task. In order to facilitate communication under challenging multipath and rapid phase variation conditions, more advanced signal processing techniques are of great interest. Incorporating of deep belief network into the UA communication system to combat the signal distortion caused by the Doppler effect and multi-path propagation [2] and UA communication channel modeling using deep learning [3] are the commonly used approaches.

Orthogonal frequency-division multiplexing (OFDM) is an effective technique to combat multipath fading in UA communication channel. In [4], an OFDM-based receiver has been developed by integrating a neural network for UA communication. However, the performance of the approach in [4] is poor in terms of recovered bits when Doppler shift exists in the system. The deep neural network (DNN) used in [4] is composed of an input layer, two fully connected layers and a regression layer as the output layer. The concept of deep learning originated from the study on artificial neural networks (ANNs). DNN consists of many layers, where each layer has a number of neurons. Each layer performs a weighted sum of the inputs followed by a nonlinear activation and the output is fed as an input to the next layer [5].

In this paper, a UA OFDM system is proposed by integrating a regression-based deep learning technique with a long short-term memory (LSTM) layer in the network architecture. An LSTM model has an artificial recurrent neural network (RNN) architecture. Its feedback connections process single data points and an entire sequences of data. The operations within the cells allow the LSTM to memorize or forget information by enabling back-propagation of the error through time and layers [6].

When the DNN is implemented in a UA communication system, the Doppler shift is compensated by the knowledge acquired from training data, and the transmitted data can be recovered without using conventional demodulation techniques. In conventional systems, Doppler shift estimation is performed by using null subcarriers [7].

By training the DNN using Doppler shifted data, the frequency offset can be learned. The DNN parameters including weights and biases of each layer are determined with respect to the training data and the DNN predicts the transmitted data with the learned channel information and frequency offset [8]. The simulation results show that the performance of regression based DNN with an LSTM layer is better than the traditional least-squares (LS) method, existing DNN with fully connected layers and convolutional neural network (CNN) in terms of bit-error-rate (BER).

II. System Model

In this paper, we consider a frame-based UA OFDM communication system. Each OFDM frame contains a pilot block and a data block as shown Fig. 1. In the data block, a binary source bit stream is mapped into data symbols drawn from the quadrature phase-shift keying (QPSK) constellation,

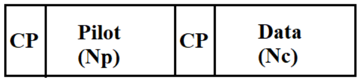

Fig. 1. frame structure.

$\boldsymbol{d} = (d[1], ..., d[N_c])^T$, where N_c is the number of data subcarriers. The pilot block contains N_p pilot subcarriers with null subcarriers at every sixth positions, $\boldsymbol{p} = (p[1], ..., p[N_p])^T$.

Each OFDM symbol is converted to the time domain by the inverse fast Fourier transform (IFFT), where a cyclic prefix (CP) with a length T_{cp} longer than the channel delay spread is added to the time domain symbol. The received signal can be written as

$$y(t) = x(t) * h(t) + w(t) \tag{1}$$

where $x(t)$ is the transmitted signal, * denotes the convolution operation, $h(t)$ is the impulse response of the channel, and $w(t)$ is the additive noise. At the receiver end, the received data frame is the input of the DNN after downshifting and removing the CP. The proposed LSTM based DNN predicts the transmitted data from $y(t)$ without explicit Doppler compensation, channel estimation, equalization and demodulation.

Doppler effect including Doppler shift and Doppler spread is an inherent feature of the communication channel, which can cause interference and affect the quality of communication. Relative movement between the transmitter and receiver and slow underwater sound propagation generates significant Doppler effect [9].

Doppler spread is the spectral broadening or compression caused by the time rate of change of the channel and is defined as the range of frequencies over of which the received Doppler spectrum is essentially non-zero. The amount of Doppler spread depends on the ratio of bandwidth to the center frequency. Assuming the power spectral density (PSD) of the transmit signal is $D_t(f)$, the Doppler affects the received PSD as $D_r(f) = 1/a^2 D_t(f/a)$, where $a = 1+v/c$, v is the relative transmitter-receiver speed and c denotes the underwater sound propagation speed. For UA communication systems, which are broadband, the shifted frequency has big difference at the lower end of the frequency band af_l and at the higher end of the frequency band af_h. This is the reason of the Doppler spreading. If f_l and f_h are comparable in the narrow-band systems, there would be only Doppler shifting and not Doppler spreading. The Doppler shift and Doppler spread are both the linear functions of the signal frequency [11]. Doppler shift of the receiver f is given by

$$f = (v/c)f_c \tag{2}$$

where v is the relative transmitter-receiver speed, c denotes the underwater sound propagation speed and f_c is the centre frequency of the band.

The current trend of Doppler compensation is a two-step processing that consists of re-sampling and residual carrier frequency offset (CFO) compensation [4], [5]. A re-sampling ratio is associated with signal compression / expansion and the ratio is determined by the estimated Doppler scaling factor. After re-sampling the effect of Doppler spreading is mitigated, and the remaining Doppler only affects the phase of the received signal.

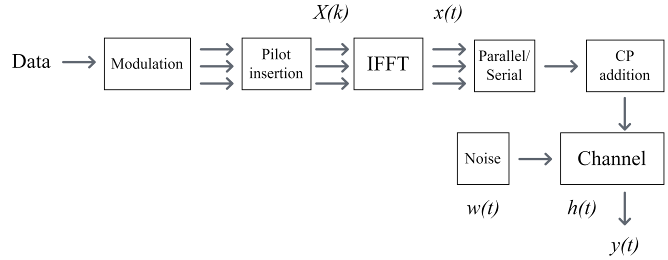

Fig. 2. Block diagram of the transmitter.

The discrete time samples of a received baseband OFDM symbol can be written as

$$\boldsymbol{y} = \boldsymbol{\Phi} \boldsymbol{F}^H \boldsymbol{D} \boldsymbol{h}_f + \boldsymbol{w} \tag{3}$$

where $\boldsymbol{y} = (y[1], ..., y[N_c])^T$ is the received data, $\boldsymbol{\Phi} = \text{diag}(e^{j2\pi f/B}, ..., e^{j2\pi N_c f/B})$, B is the bandwidth, $\boldsymbol{F}$ is the $N_c \times N_c$ discrete Fourier transform matrix, $(.)^H$ denotes the conjugate transpose,

$$\boldsymbol{D} = \begin{cases} \text{diag}(\boldsymbol{d}) \text{ for the data block} \\ \text{diag}(\boldsymbol{p}) \text{ for the pilot block} \end{cases} \tag{4}$$

diag(.) denotes a diagonal matrix and $\boldsymbol{h}_f = (h_f[1], ..., h_f[N_c])^T$ is the channel frequency response.

III. System Design

This section explains the system design for the proposed OFDM transmitter and the receiver.

1) Transmitter: Each transmitted frame contains one OFDM data block and one pilot block. In each data block, there are $N_c = 64$ subcarriers and the pilot block consists of $N_p = 64$ pilot subcarriers with null subcarriers in every sixth positions. The data symbols are modulated by the QPSK constellations. Hence, one symbol is encoded by two bits. The cyclic prefix is chosen as $T_{cp} = 20$ samples. The proposed transmitter architecture is illustrated in Fig. 2.

2) Receiver: The signals received from the transmitter are converted to the frequency domain using the fast Fourier transform (FFT) and the baseband signals are input into the trained DNN as shown in Fig. 3. The output of the DNN is the demodulated transmitted bits. Training the network is the vital part of any system which includes DNN. Training data including the transmitted data and the received data is illustrated in Fig. 4. The DNN extracts the channel information and Doppler effects from the received signal and tunes its internal parameters accordingly. The weights and biases of each layer of the DNN are determined using the stochastic gradient descent and the back-propagation algorithm. The mean-squared error L formulated in (5) is used as the loss function

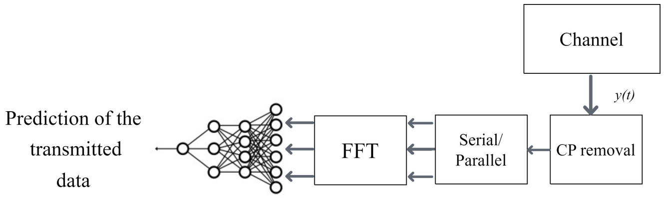

Fig. 3. Block diagram of the receiver.

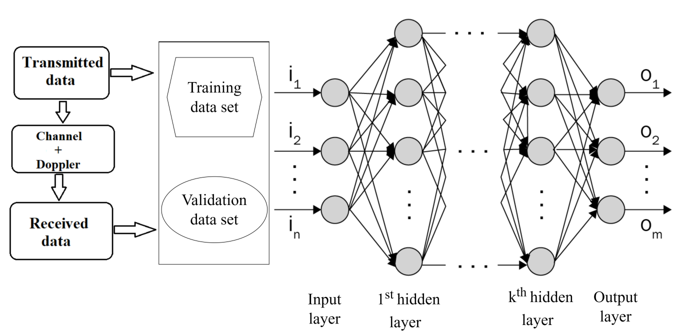

Fig. 4. Training process of the DNN.

which evaluates the difference between the transmitted data and the DNN predictions

$$L = \frac{1}{N}\sum_{k=0}^{N-1}(\hat{b}(k) - b(k))^2 \tag{5}$$

where N is the number of bits, $\hat{b}(k)$ is the predicted bit, and $b(k)$ is the training bit [4].

In the proposed system, an LSTM layer is introduced to the DNN by replacing a fully connected layer after the input layer. This LSTM based DNN consists of four layers as shown in Fig. 5, including a sequence input layer, an LSTM layer, a fully connected layer and a regression layer. The performance of the CNN and a fully connected DNN are also used to compare to that of the LSTM, where the BER formulated in (5) is used as the performance metric.

The proposed approach uses the DNN to compensate the Doppler shift and it replaces the channel estimation, equalization and demodulation in the traditional receiver. In the traditional LS method, CFO compensation is performed by minimizing the leakage energy in the null subcarriers introduced in the pilot OFDM block [12]. The compensation of the CFO on the received baseband symbol is performed by

$$d[n] = y[n]e^{-j2\pi N_c \hat{f}/B} \tag{6}$$

where $\hat{f}$ is the estimated value of CFO and is generated for each OFDM block by minimizing the energy of the null subcarriers. The cost function is defined as

$$J(f) = \sum_{k\in S_N} |\boldsymbol{f}_k^H \boldsymbol{\Phi}^H(f)\boldsymbol{y}|^2 \tag{7}$$

where, S_N is the set of null subcarriers and $\boldsymbol{f}_k$ is

$$\boldsymbol{f}_k = [1, e^{j2\pi k/N_c}, ..., e^{j2\pi k(N_c-1)/N_c}]^T \tag{8}$$

The estimate of f is given by

$$\hat{f} = \arg\min_f J(f) \tag{9}$$

A. LSTM

Recurrent neural networks with an LSTM layer have emerged as an effective and reliable model for several learning problems related to sequential data. A memory cell retains its state over time and nonlinear gating units control the information flow into and out of the cell which creates the core of the LSTM architecture as in Fig. 7 [13]. In Fig. 6, the peephole connections are illustrated as the blue connections from the cell to the gates, and the forget gate are added to the architecture in order to make precise timings easier to learn. The forget gate enables the LSTM to reset its own state, Input gate decides which new information is going to enter the state of LSTM. The output gate updates and finalizes the next hidden state [14].

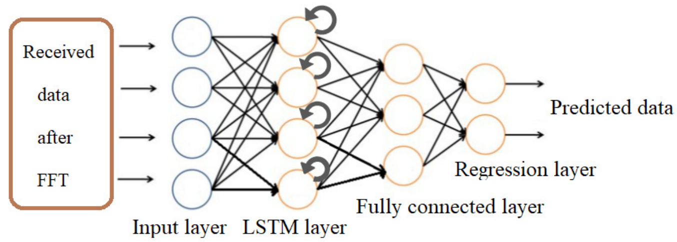

Fig. 5. DNN with LSTM architecture.

B. CNN

A CNN is composed of several layers, including an image input layer, a convolutional layer and a fully connected layer. A regression layer is chosen as the output layer. The convolution layer extracts the feature representations of the inputs. As

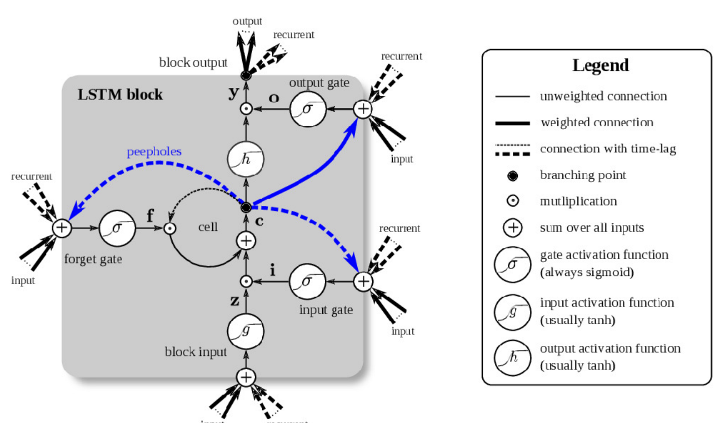

Fig. 6. LSTM memory cell [14].

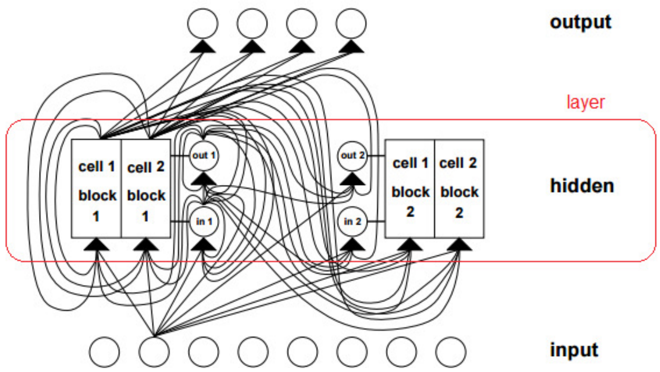

Fig. 7. LSTM layer architecture [13].

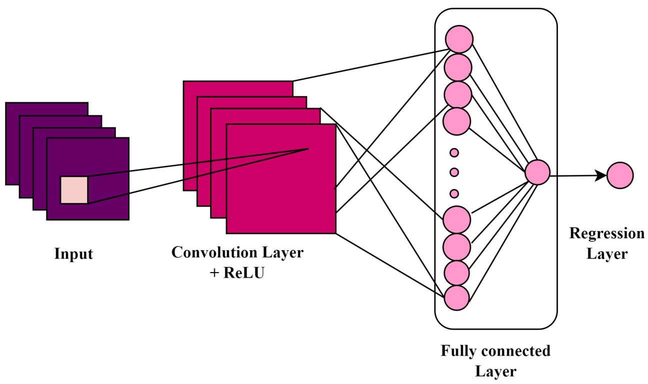

Fig. 8. CNN architecture.

TABLE I
NETWORK CONFIGURATION AND DETAILS.

Network	Layers	Neurons	Learnables	Total learnables
LSTM	Sequence input layer	40		0
	LSTM layer	80	Input weights: 320x40 Recurrent weights: 320x80 Bias : 320x1	38720
	Fully connected layer	4	Weights: 4x80 Bias:4x1	324
	Regression layer			0
FC	Sequence input layer	40		0
	Fully connected layer	80	Weights:80x40 Bias:80x1	3280
	ReLU	80		0
	Fully connected layer	4	Weights:4x80 Bias:4x1	324
	Regression layer			0
CNN	Image input layer	20x2 x1		0
	Convolution layer	20x2x8	Weights: 4x1x1x8 Bias: 1x1x8	40
	ReLU	20x2x8		0
	Fully connected layer	1x1x4	Weights: 4x320 Bias: 4x1	1284
	Regression layer			0

shown in Fig. 8, a convolution layer is composed of several convolution kernels which are used to compute different feature elements of the input. Although these kernels usually have very low spatial dimensions, they unfold over the entire input depth dimension. A convolution layer convolves each filter across the input data and produces an activation map, once the information reaches it [15]. In this simulation, two dimensional input data is fed to the image input layer. The convolution layer filters extract the specific features from input data and feed the processed data to the fully connected layer. A regression layer is used as the output layer, which receives the information from the fully connected layer and produces the output.

IV. EXPERIMENT RESULTS

In this section, we evaluate the performance of our proposed UA OFDM system by comparing the performance of DNN with an LSTM layer, DNN with fully connected layer, CNN and the conventional LS based receiver. The proposed DNN architecture is shown in Fig. 5. CNN architecture is shown in Fig. 8. We simulate a UA channel with 15 paths and the channel delay spread is 15 samples.

Normally, training the network with a large amount of data and high number of layers and neurons can produce better performance. However, long training time is required. When the trained network is used in a communication system which includes a time varying channel, it will not produce accurate predictions since the newly fed data have a huge difference from the training data set. Since the channel is time varying, the DNN needs to adapt to the changing channel. The original trained parameters need to be retrained. Small number of parameters requires less training time and suits for real time implementation. To achieve this, we need to use small training data set which requires less number of parameters for the network. By selecting only 10 data subcarriers and 10 pilot subcarriers for the feature extraction process during the training data generation allows us to reduce the parameters of the networks and leads to a short training time.

Table I shows the network configuration of each DNN and its learnable parameters used in the simulation, where 500 OFDM packets of 10 data subcarries and 10 pilot subcarriers are used for training data generation from 64 data subcarriers and 64 pilot subcarriers. 4 bits are recovered per OFDM frame in the validation. The DNN with LSTM layer consists of 40 neurons in the input layer, 80 neurons in the LSTM layer and 4 neurons in the fully connected layer. The DNN with fully connected layers has the same number of neurons in each layer as the DNN with LSTM layer. For the CNN, the size of image input layer is 20-by-2, convolution layer has 8 filters with the size of 4-by-1 and a fully connected layer has 4 neurons. The number of neurons for the input layer is twice the

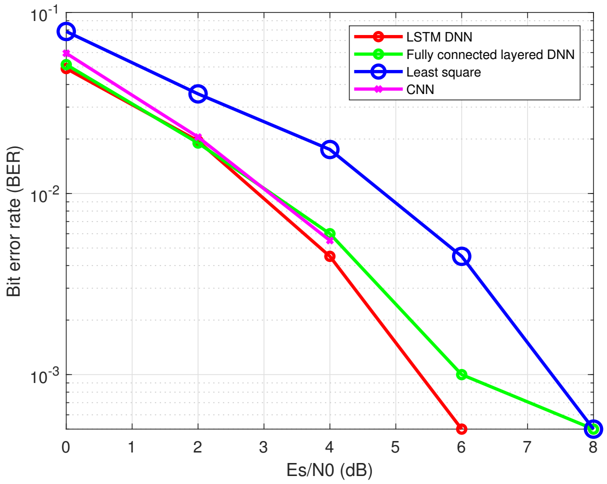

Fig. 9. System BER performance; $v = 1m/s$ for both training (500 OFDM packets) and testing.

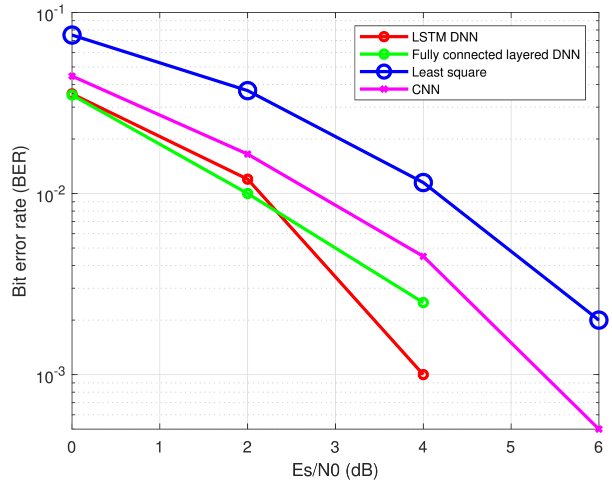

Fig. 10. System BER performance; $v = 1m/s$ for both training (1000 OFDM packets) and testing (4 bits recovered per frame).

number of subcarriers chosen for training data generation for both DNN with LSTM layer and DNN with fully connected layers, as complex QPSK constellations are used. The number of neurons for the fully connected layer in front of the output layer is the number of bits recovered per OFDM frame.

Firstly, we consider the scenario where the same relative speed $(v = 1m/s)$ is used to train and validate the DNNs. The system performance is illustrated in Fig. 9, which shows that the BER performance of the DNN with LSTM system is much better than the results obtained by the LS method, the CNN and the fully connected layered DNN without LSTM. In this scenario, only 500 OFDM packets are used to generate the training data sets, which includes, 400 training data sets and 100 validation sets. As the number of training packets increases the performance of each network is improved which is shown in Fig. 10, where 1000 OFDM packets are used. Table I illustrates that, the number of learnable parameters is directly related to the performance of each network. As the number of internal parameters increases the network learns more information from the training data and improves its performance.

1000 OFDM packets are utilized in the generation of training data sets for Figs. 10 - 13, where the same relative speed $(v = 1m/s)$ is used to train and validate the DNNs. For testing, 4, 6, 12 and 20 bits are recovered respectively, which are the dimensions of the labels and network input/output sizes. For example, if we are aiming to recover 10 bits, then we need to provide 10 bits to the network as labels during the training process. These figures show that, DNN with LSTM performs better than other two networks and the LS method.

Secondly, we validate the DNN performance with different Doppler shifts for training and testing, where the DNN with an LSTM layer and other two networks are trained with $v = 1m/s$ and tested with $v = 1.2m/s$. These network performances are compared with the LS method with the

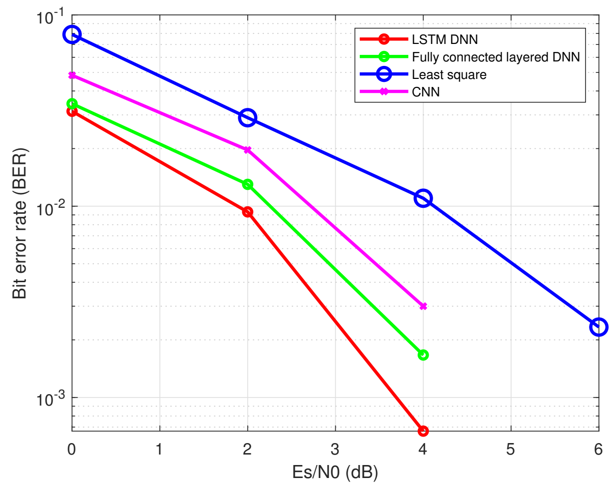

Fig. 11. System BER performance; $v = 1m/s$ for both training (1000 OFDM packets) and testing (6 bits recovered per frame).

same Doppler shift $(v = 1.2m/s)$. Figs. 14 and 15 show that better performance can be achieved by the DNN with LSTM layer system compared with the method in [4], CNN and the LS method. Surprisingly, the neural networks are able to perform better than the traditional LS method although they are trained for a specific frequency offset value but tested with different frequency offset values. Also, the results show that the BER performance of the DNN and CNN systems can be improved by increasing the numbers of subcarriers and numbers of OFDM packets used for training the networks, increasing numbers of neurons in individual layers, changing the epochs rate and other parameters of the neural network. In this simulation, the networks are trained with mini-batch size of 50 and 100 epochs. As illustrated in Fig. 9 and Fig. 10, when the number of OFDM packets increased form 500 to 1000, the performance of networks improves.

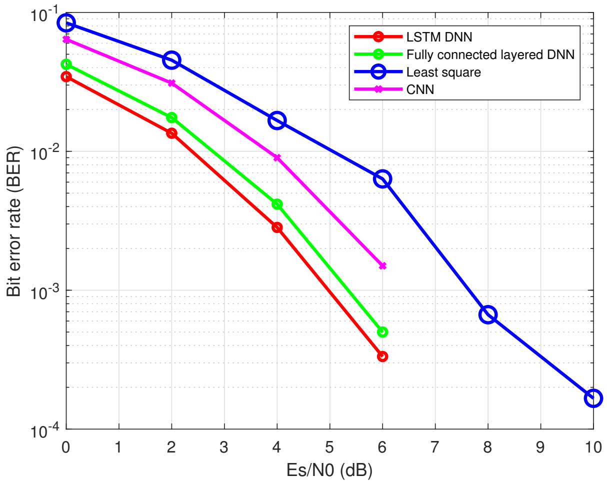

Fig. 12. System BER performance; $v = 1m/s$ for both training (1000 OFDM packets) and testing (12 bits recovered per frame).

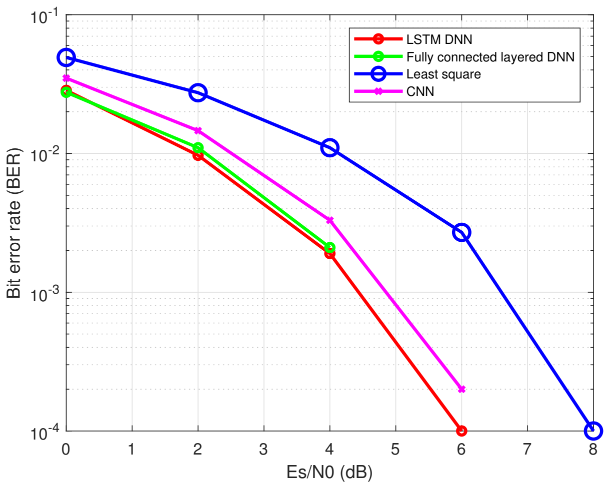

Fig. 13. System BER performance; $v = 1m/s$ for both training (1000 OFDM packets) and testing (20 bits recovered per frame).

V. Conclusion

In this paper, a UA OFDM receiver integrated with an LSTM-based DNN is proposed to mitigate the Doppler effect in UA system. The results show that the proposed approach achieves better performance compared with the traditional receiver which uses the LS estimation method, DNN with fully connected layer and CNN in Doppler compensation, demodulating and decoding the transmitted bits. Even the training data and testing data vary in the frequency offset value, the neural networks perform better than the conventional method. In the future, the performance of the proposed receiver can be verified by using real time UA communication for training the DNN and validating the performance.

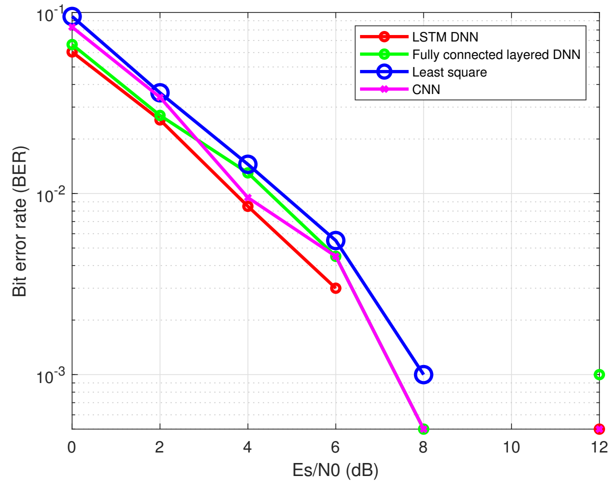

Fig. 14. System BER performance; different $v = 1.2m/s$ values for training and testing (4 bits recovered per frame).

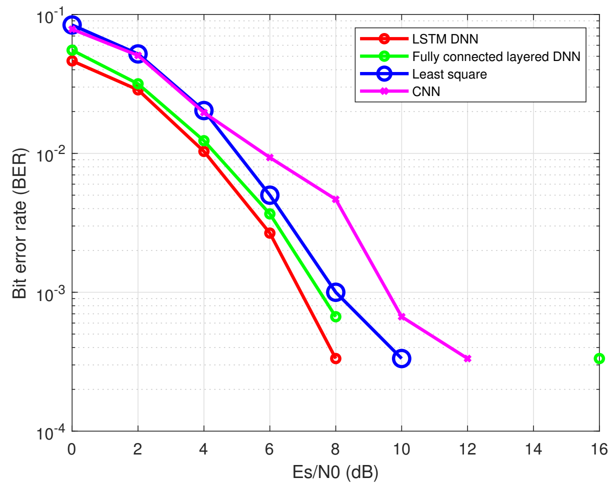

Fig. 15. System BER performance; different $v = 1.2m/s$ values for training and testing (6 bits recovered per frame).

References

[1] P. Qarabaqi and M. Stojanovic, "Statistical characterization and computationally efficient modeling of a class of underwater acoustic communication channels," *IEEE Journal of Oceanic Engineering*, vol. 38, no. 4, pp. 701-717, 2013.

[2] L. L. Abigail, Y. Chau and H. Dorien, "Underwater acoustic communication receiver using deep belief network," *IEEE Transactions on Communications*, vol. 69, no. 6, pp. 3698-3708, 2021.

[3] O. Oluwaseyi, A. Damilola and Q. Lijun, "Underwater acoustic communication channel modeling using deep learning," *Association for Computing Machinery*, no. 9, pp. 8, 2022. url = https://doi.org/10.1145/3491315.3491323.

[4] Y. Zhang, J. Li, Y. Zakharov, X. Li, and J. Li, "Deep learning based underwater acoustic OFDM communications," *Applied Acoustics*, vol. 154, pp. 53-58, 2019.

[5] M. Grégoire, S. Wojciech and M. Klaus-Robert, "Methods for interpreting and understanding deep neural networks," *Digital Signal Processing*, vol. 73, pp. 1-15, 2018.

[6] A. Graves, "Long short-term memory," *Supervised Sequence Labelling with Recurrent Neural Networks*, pp. 37-45, 2012.

[7] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," *IEEE Journal of Oceanic Engineering*, vol. 33, no. 2, pp. 198–209, Apr. 2008.

[8] H. Ye, G. Y. Li, and B.-H. 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, 2017.

[9] P. Chen, Y. Rong, S. Nordholm, Z. He, and A. J. Duncan, "Joint channel estimation and impulsive noise mitigation in underwater acoustic OFDM communication systems," *IEEE Transactions on Wireless Communications*, vol. 16, no. 9, pp. 6165-6178, 2017.

[10] Z. Xiao, H. Xiao, Y. Jingwei and S. Xue-li, "Study on Doppler effects estimate in underwater acoustic communication," *Proceedings of Meetings on Acoustics*, vol. 19, no. 1, pp. 3698-3708 url = https://asa.scitation.org/doi/abs/10.1121/1.4799512,

[11] B.S. Sharif, J. Neasham, O.R. Hinton, and A.E. Adams, "A computationally efficient Doppler compensation system for underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 25, no. 1, pp. 52-61, 2000.

[12] B, Reza, "Impulsive noise detection and mitigation in communication systems." *Kansas State University*, 2019.

[13] H. Sepp and S. Jürgen, "Long short-term memory," *Neural Computation*, vol. 9, pp. 1735–1780, 1997.

[14] G. Klaus, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A search space odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222-2232, 2017.

[15] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, B. Shuai, T. Liu, X. Wang, G. Wang, J. Cai and T. Chen, "Recent advances in convolutional neural networks," *Pattern Recognition*, vol. 77, pp. 354-377, 2018, url = https://www.sciencedirect.com/science/article/pii/S0031320317304120.