Benefits of Neural Network-based Receivers in Underwater Acoustic Communication Systems

Sabna Hassan, Peng Chen, Yue Rong, Kit Yan Chan
Curtin University

Iftekhar Ahmad
Edith Cowan University

Jinhong Yuan
University of New South Wales

***Abstract*—In conventional orthogonal frequency-division multiplexing (OFDM) communication systems, channel knowledge on data subcarriers is obtained through linear interpolation of the estimated channel information on pilot subcarriers. This operation degrades the accuracy of channel estimation, particularly when the multipath channel has a large delay spread as in underwater acoustic (UA) channels. In this paper, we show that such drawback can be overcome by a neural network (NN)-based receiver, which is effective in learning the nonlinearity in the channel information on subcarriers. A multilayer perceptron (MLP) network, a convolutional neural network (CNN), and a long short-term memory (LSTM) network architectures are attempted. These NNs are trained and tested by data collected in a real-world UA communication experiment conducted in a water tank. The results show that the MLP and LSTM network-based receivers achieve better bit-error-rate (BER) performance than the conventional OFDM receiver.**

I. Introduction

Underwater acoustic (UA) communication through a multipath channel creates multiple arrivals from various paths [1]-[3]. Orthogonal frequency-division multiplexing (OFDM) communication is applied to mitigate the multipath interference [4].

Conventional OFDM receivers use linear interpolation for channel estimation [4]. However, in UA channels, the channel frequency response often presents strong nonlinearity between two pilot subcarriers, due to the large delay spread of the multipath channel. As the channel delay profile is unknown in practice, this nonlinearity cannot be modeled precisely. Such unknown nonlinearity motivates the neural network (NN)-based receiver design, which is effective in learning the nonlinearity through the NN training. The NN consists of a number of independent layers, where each layer has a certain number of nodes. In general, 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]. Through the training process, the weights and biases are tuned according to the provided training data. A loss function is used for the parameter estimation of a supervised deep learning. During the training the loss function is used to optimize the optimal weights and biases [6].

In this paper, a UA OFDM system is proposed by integrating a regression-based NN receiver. Three commonly used NN architectures are investigated: a multilayer perceptron (MLP) network, a convolutional neural network (CNN), and a long short-term memory (LSTM) network. The MLP network consists of an input layer, fully connected hidden layers and a regression output layer [7]. A CNN is an MLP with at least one convolutional hidden layer. Convolutional layer includes multiple optimizable filters and the number of filters defines the depth of a convolutional layer [8]. An LSTM model has a gated recurrent neural network (RNN) architecture, whose feedback connections can process single data points and an entire sequences of data. The operations within the cells allow the LSTM to keep or forget the information by enabling back-propagation of the error through the time and the layers hence preserve them [9].

The NN-based receivers are trained and tested through data collected in a recent UA communication experiment conducted in a water tank. The transmitted data can be recovered without using conventional demodulation techniques. The NN parameters including weights and biases of each layer are determined according to the training data and the NN predicts the transmitted data with the channel information learned from the training data [10]. The experimental results show that for various signal-to-noise ratios (SNRs), the bit-error-rate (BER) performance of the MLP and LSTM NNs is better than the conventional UA OFDM system, which uses a least-squares (LS)-based channel estimator.

The reminder of this paper is as follows. In Section II, benefits of introducing the NN in the UA OFDM system are presented. The proposed system design is explained in Section III and an overview of the training process and architecture of various NNs are given in Section IV. Experimental results are explained in detail in Section V. Finally, conclusions and future works are drawn in Section VI.

II. Benefits of NN-based OFDM Receiver

We consider a frame-based UA OFDM communication system. Each OFDM frame contains a pilot block and a data block. In the data block, a binary source bit stream is mapped into data symbols drawn from the quadrature phase-shift keying (QPSK) constellation $\boldsymbol{d} = (d[1], ..., d[N_c])^T$, where N_c is the number of data subcarriers. The pilot block contains N_p pilot subcarriers $\boldsymbol{p} = (p[1], ..., p[N_p])^T$ with null subcarriers at each of the sixth position. Each OFDM symbol is converted to the time domain by the inverse fast Fourier transform (IFFT), and a cyclic prefix (CP) with the 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 channel impulse response, and $w(t)$ is the additive noise. At the receiver end, the received data frame is input to the NN after downshifting and removing the CP. The NN directly predicts the transmitted data from $y(t)$ without explicit channel estimation, equalization and demodulation.

To illustrate the limitation of the conventional OFDM receiver, let us consider a multipath UA channel whose frequency response at the kth subcarrier is given by

$$H_k = \sum_{l=1}^{L} h_l e^{-j2\pi f_k \tau_l}, \quad k = 1, \ldots, N_c \tag{2}$$

where h_l and τ_l, $l = 1, \ldots, L$, are the amplitude and delay of the lth arrival, respectively, $f_k = k/T$ is the frequency of the kth subcarrier, and T is the OFDM symbol duration. In conventional OFDM systems with comb-based pilot pattern, pilot subcarriers are used for channel estimation. The channel information on a data subcarrier is obtained by performing linear interpolation of the estimated channel coefficients of two most adjacent pilot subcarriers. This technique performs well in terrestrial radio systems, since the multipath spread is much shorter than the length of an OFDM symbol (i.e. $\tau_l \ll T$), However, the performance of this technique may degrade in UA communication systems, since UA channels usually have a large delay spread and later arrivals are possible to have larger amplitudes than early arrivals. In these scenarios, the channel frequency response (2) presents strong nonlinearity between two pilot subcarriers. Such unknown nonlinearity can be effectively learned by a NN-based receiver.

The Doppler shift compensation in conventional method is performed by minimizing the leakage energy in the null subcarriers introduced in the pilot OFDM block [11]. The compensation of the carrier frequency offset (CFO) on the received baseband symbol is performed by

$$d[n] = y[n] e^{-j2\pi n \hat{f}/B_w} \tag{3}$$

where $y[n]$ is the received signal sample, $d[n]$ is the CFO-compensated sample, $\hat{f}$ is the estimated value of the CFO, and B_w is the bandwidth. The CFO is estimated for each OFDM block by minimizing the energy of the null subcarriers. The objective function of the CFO estimation is defined as

$$J(f) = \sum_{k \in S_N} |\boldsymbol{f}_k^H \boldsymbol{\Phi}^H(f) \boldsymbol{y}|^2 \tag{4}$$

where $\boldsymbol{y} = (y[1], ..., y[N_c])^T$ is the received data, S_N is the set of null subcarriers, $\boldsymbol{f}_k$ and $\boldsymbol{\Phi}$ are

$$\boldsymbol{f}_k = [\, 1, e^{j2\pi k/N_c}, ..., e^{j2\pi k(N_c-1)/N_c}\,]^T \tag{5}$$

and

$$\boldsymbol{\Phi} = \mathrm{diag}(e^{j2\pi f/B_w}, ..., e^{j2\pi N_c f/B_w}). \tag{6}$$

Here $(.)^H$ denotes the conjugate transpose and $\mathrm{diag}(.)$ denotes a diagonal matrix. The estimate of f is given by

$$\hat{f} = \arg\min_f J(f) \tag{7}$$

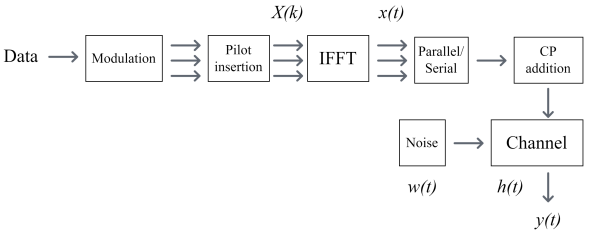

Fig. 1. Block diagram of the transmitter.

By training the NN using the Doppler shifted data, the frequency offset can be learned without explicit estimation required by the conventional methods. The NN parameters including weights and biases of each layer are determined with respect to the training data and the NN predicts the transmitted data with the learned channel information and the frequency offset [12].

III. System Design

1) Transmitter: Each transmitted frame contains two OFDM blocks, each having $N_c = 107$ subcarriers. The first block consists of 90 pilot subcarriers and 17 evenly distributed null subcarriers. All the subcarriers in the second OFDM block are data subcarriers. The data symbols are modulated by the QPSK constellations. Hence, one symbol is encoded by two bits. The CP is chosen as $T_{cp} = 25$ ms and the system bandwidth is $B_w = 4$ kHz. The transmitter block diagram is illustrated in Fig. 1.

2) Receiver: The received signals are converted to the frequency domain using the fast Fourier transform (FFT) and the baseband signals are input into the NN. The output of the NN is the demodulated transmitted bits as shown in Fig. 2. The NN is trained by the data collected from a recent UA communication experiment in a tank. The weights and biases of each layer of the NN are determined using the stochastic gradient descent and the back-propagation algorithm. The mean-squared error L formulated in (8) is adopted as the loss function, which evaluates the difference between the transmitted data and the NN predictions.

$$L = \frac{1}{N} \sum_{k=0}^{N-1} (\hat{b}(k) - b(k))^2 \tag{8}$$

where N is the number of bits, $\hat{b}(k)$ is the predicted bit, and $b(k)$ is the training bit. The proposed approach uses NN to replace the channel estimation, equalization, and demodulation in the conventional OFDM receiver.

IV. Training the NNs

The performance of a NN depends on the quality and quantity of the training data provided. The depth of a NN used for a particular problem is determined with respect to the size of the training data. Small data set may overfit the NN with a large number of layers while the same NN trained with a large data set could generalize well. Even though, a large data

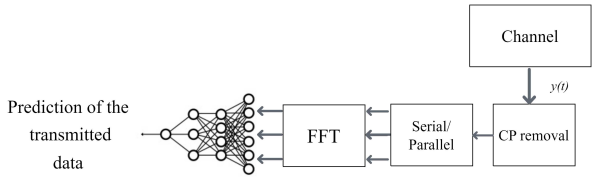

Fig. 2. Block diagram of the receiver.

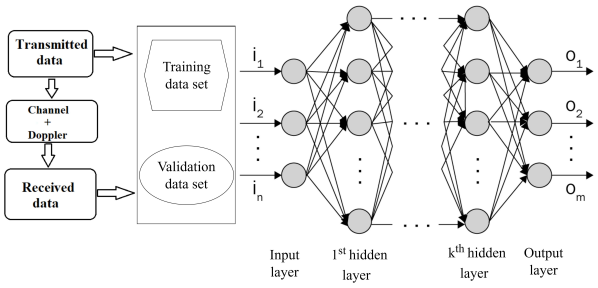

Fig. 3. NN training process.

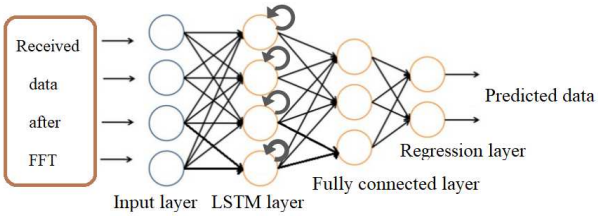

Fig. 4. Architecture of the LSTM-based receiver.

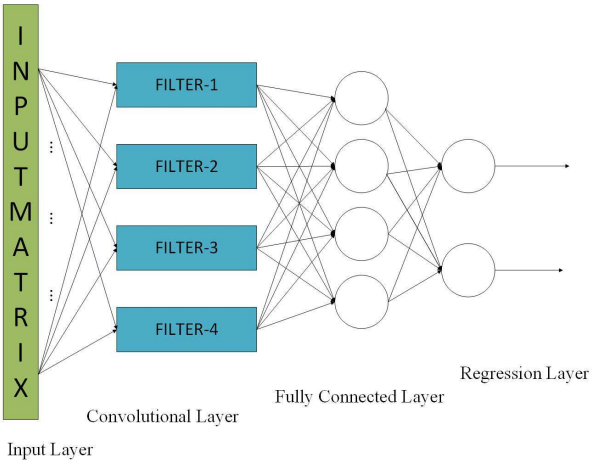

Fig. 5. Architecture of the CNN-based receiver.

set gives an opportunity to increase the number of layers of the network, it is better to start the network with a minimum number of layers. The number of layers, neurons, epochs, batch size, learning rate, etc required for the suitable network for a particular problem is decided through trial and error method. To generate training data for the NN in the proposed supervised learning method, both the transmitted and the received data are used as shown in Fig. 3. From this training data set, the NN learns the relationship between transmitted and received data and the channel characteristics and tune its internal parameters accordingly. Half of the recorded data from the receiver is used for training and validation purpose. The other half is used as testing data for the final evaluation. Alternative packets of the received data are selected for the training data as it gives the whole channel information to the NN.

A. MLP

MLP is the most commonly used NN with an input layer, an output layer and at least one hidden layer as shown in Fig. 3. It has a feed-forward architecture as there is no internal loop and the output of one neuron does not affect itself. The non-linear activation functions, loss functions and optimizers are the functions used to guide the performance of the MLP [13]. These functions, the number of layers, neurons and other parameters are decided through trial and error process. The MLP used in the proposed design has four layers: a sequence input layer as the input layer, two fully connected layer as the hidden layers and a regression output layer.

B. LSTM

For sequential data related problems LSTM is the suitable network to work with. It is an MLP with an LSTM layer after the input layer as shown in Fig. 4. Memory cells in the LSTM layer provide the capability of reset or retain the current state of the model [14]. The architecture of this NN is a sequence input layer followed by an LSTM layer, a fully connected layer and a regression output layer.

C. CNN

CNN is commonly used in image processing related problems. CNN is a multilayered NN with an image input layer, a convolutional layer and a fully connected layer as hidden layers and a regression output layer as shown in Fig. 5. The convolutional layer is composed of a number of filter weights also known as kernels and the kernels extract the feature compositions of the data given [15]. In this sequential data problem, a matrix with two column data is fed to the image input layer and four filters with size [4, 4] is used in the convolutional layer.

V. Experimental Results

In this section, we compare the performance of the proposed NN-based receivers with the conventional OFDM receiver. We transmitted 2000 OFDM frames six times with the transmitter gain of -10dB, -14dB, -18dB, -22dB, -26dB, and -30dB respectively. By varying the transmitter gain, the SNR at the receiver changes. To train the NNs, only -10dB gain is used, but for testing, all the six gains are used. Fig. 6 shows the channel profile of the water tank environment illustrated in Fig. 7 with length, width, and depth as 2.5 m, 1.5 m, and 1.8 m respectively. The transmitter and receiver are 3 m apart

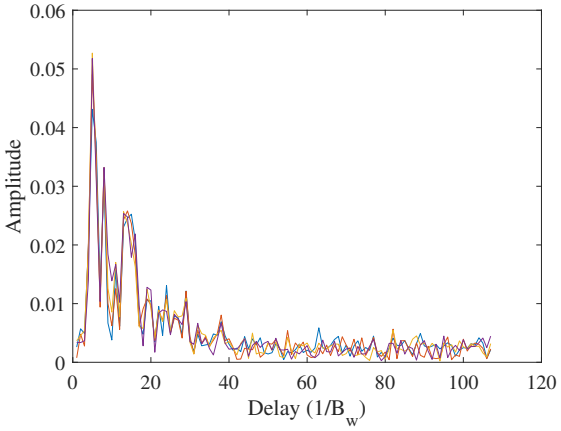

Fig. 6. Multipath channel profile.

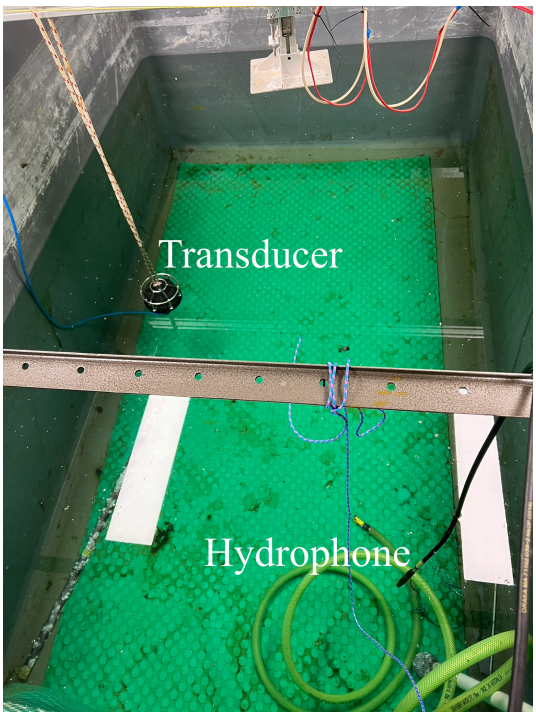

Fig. 7. Tank set up.

during the experiment. We can see that the multipath delay spread has a length of half OFDM symbol duration, which is much longer than that of terrestrial radio systems.

The performance of the proposed system is tested in various scenarios as shown follows.

1) Fixed water level: The system performance when the receiver is fixed at a location is illustrated in Fig. 8, which shows that the BER performance of the LSTM-based and MLP-based receivers is better than those of the conventional OFDM receiver and the CNN-based receiver. Note that the NNs are trained at the -10dB gain, but they show good performance at all six gains tested, indicating that the NNs generalize well.

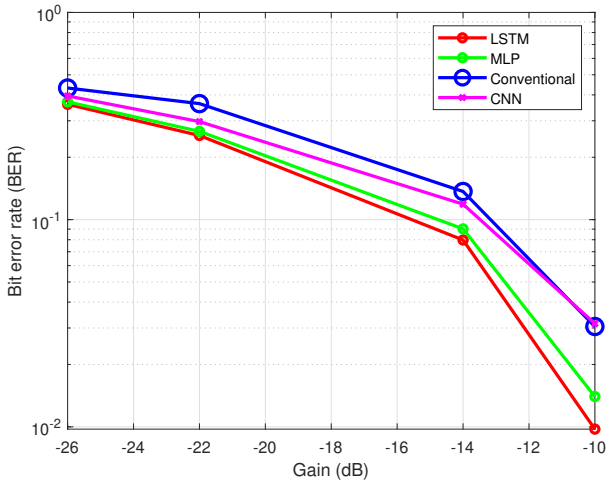

Fig. 8. System BER performance for fixed water level data.

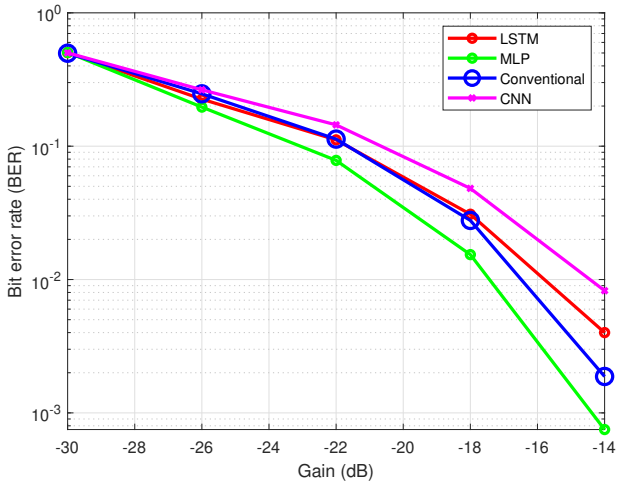

Fig. 9. Trained with mix data and tested with high water level data; 4 bits recovered.

2) Mixed water level: Fig. 9 to Fig. 11 show the performance when the NNs are trained with a mix of received data in high water level and low water level, and tested with high water level data. Fig. 12 to Fig. 14 illustrate the performance when the NNs are trained with the mixed data, but tested with the low water level data set. These two data sets are obtained by placing the receiver in two different depths of the water tank. The testing data is not included in the training data. The result shows that the MLP performs better than the traditional method.

3) Varying subcarriers and epochs: We also find that the BER performance of the NN system can be improved by varying the number of subcarriers used for training and testing. The increase or decrease in bits recovered affects the performance of both NNs and the conventional method, which is shown in Fig. 9 and Fig. 10. Changing the epochs rate also impact the performance of the NNs, as large epochs rate

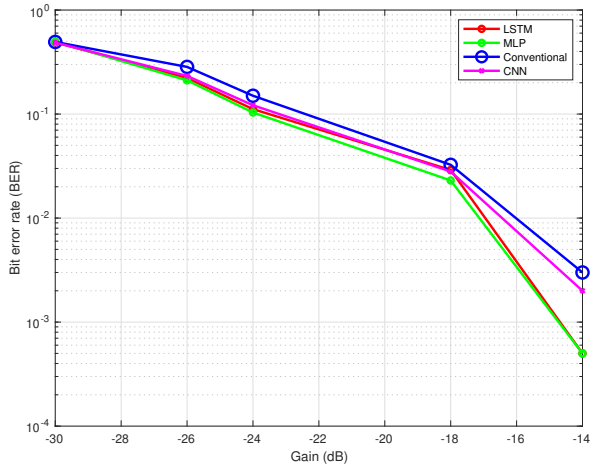

Fig. 10. Trained with mix data and tested with high water level data; 2 bits recovered.

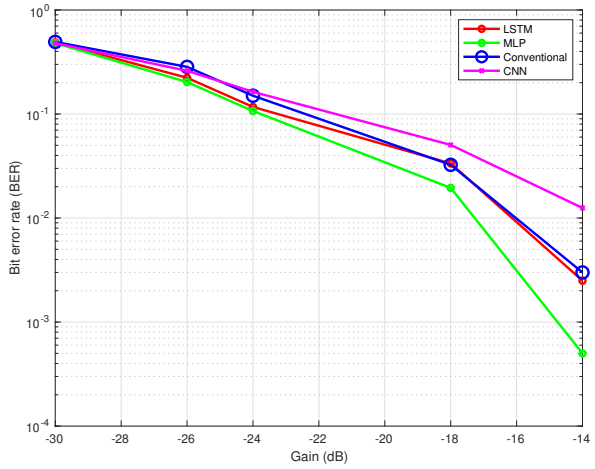

Fig. 11. 2 bits recovered with less number of epochs.

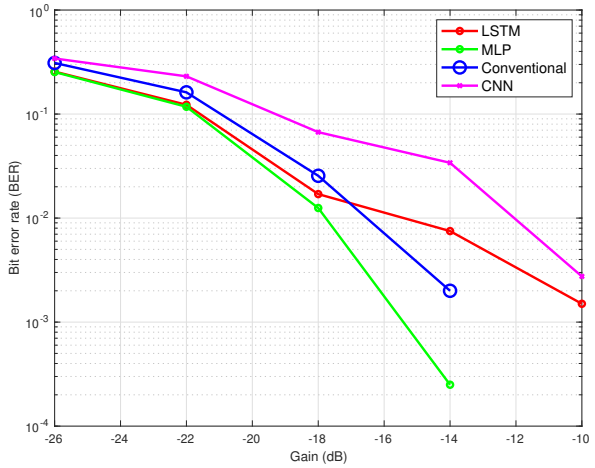

Fig. 12. Trained with mix data and tested with low water level data; 1000 packets.

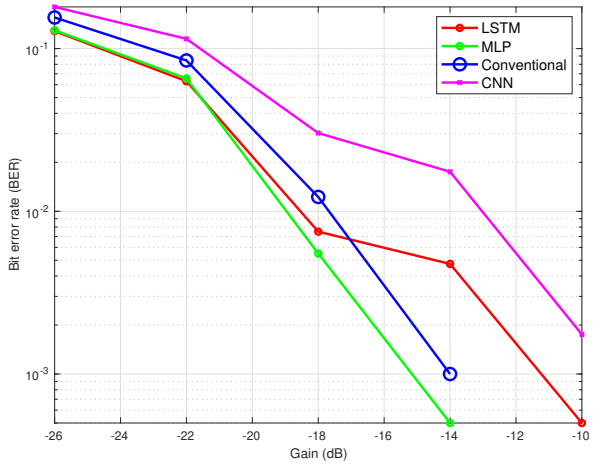

Fig. 13. Trained with mix data and tested with low water level data; 500 packets.

provides more opportunities to extract the characteristics of the training data, and improves the internal parameters tuning. This is evident in Fig. 10 with 1500 epochs used during training, whereas only 100 epochs are used in Fig. 11.

4) Varying the test data size: Fig. 12, Fig. 13, and Fig. 14 show the difference in performance for various number of testing data packets: 1000, 500, and 250 respectively by keeping the size of training data set the same. A mix of high water level data and low water level data is used as the training data and low water level data is used for testing in these results.

VI. Conclusions

We demonstrated the benefits of using NN-based receivers in UA OFDM systems. Through NN training, these receivers effectively learn the nonlinearity in the channel frequency response between pilot subcarriers. Tank trial results show that the proposed approach achieves better performance compared with the conventional UA OFDM receiver, which uses linear interpolation for channel estimation. In the future, the performance of the system will be evaluated in a river test and more investigations are planned in the scenario of moving the transmitter or the receiver.

References

[1] F. De-Rango, F. Veltri, and P. Fazio, "A multipath fading channel model for underwater shallow acoustic communications," in *Proc. IEEE ICC*, Ottawa, ON, Canada, pp. 3811–3815, 2012.

[2] J. Shao, C. Zhang, Y. Xie, D. Mishra, Y. Rong, P. Chen, and J. Yuan, "Fractional Fourier transform based channel estimation in underwater acoustic communications," in *Proc. MTS/IEEE OCEANS*, Singapore, Apr. 14-18, 2024.

[3] C. Zhang, Y. Xie, D. Mishra, T. Pacino, J. Shao, B. Li, J. Yuan, P. Chen, and Y. Rong, "A low complexity channel emulator for underwater acoustic communications," in *Proc. MTS/IEEE OCEANS*, Gulf Coast, Mississippi, USA, Sep. 25-28, 2023.

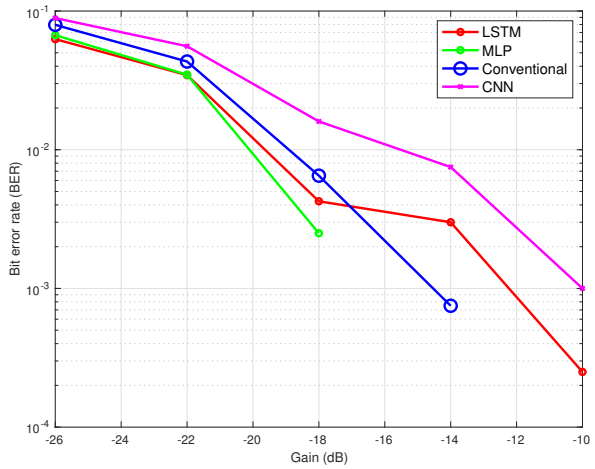

Fig. 14. Trained with mix data and tested with low water level data; 250 packets.

[4] P. Chen, Y. Rong, S. Nordholm, and Z. He, "An underwater acoustic OFDM system based on NI compactDAQ and LabVIEW," *IEEE Systems Journal*, vol. 13, pp. 3858–3868, Dec. 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] B. Vasileios, R. Christian, C. Gustavo, and N. Nassir, "Robust optimization for deep regression," *IEEE International Conference on Computer Vision (ICCV)*, 2015.

[7] 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.

[8] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on convolutional neural networks (CNN) in vegetation remote sensing," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 173, pp. 24-49, 2021.

[9] S. Lathuilière, P. Mesejo, X. Alameda-Pineda, and R. Horaud, "A comprehensive analysis of deep regression," *IEEE Trans Pattern Analysis Machine Intelligence*, vol. 42, no. 9, pp. 2065–2081, 2019.

[10] 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.

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

[12] S. Hassan, P. Chen, Y. Rong, and K. Y. Chan, "Doppler shift compensation using an LSTM-based deep neural network in underwater acoustic communication systems," *IEEE OCEANS 2023-Limerick*. pp. 1-7, 2023.

[13] P. Marius-Constantin, V. E. Balas, L. Perescu-Popescu, and N. Mastorakis, "Multilayer perceptron and neural networks," *WSEAS Transactions on Circuits and Systems*, vol. 8, no. 7, pp. 579-588, 2009.

[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.