LETTER

An Optimized CNN-Attention Network for Clipped OFDM Receiver of Underwater Acoustic Communications

Feng LIU[†a)], **Qian XI**[†], *Nonmembers*, *and* **Yanli XU**[†], *Member*

SUMMARY In underwater acoustic communication systems based on orthogonal frequency division multiplexing (OFDM), taking clipping to reduce the peak-to-average power ratio leads to nonlinear distortion of the signal, making the receiver unable to recover the faded signal accurately. In this letter, an Aquila optimizer-based convolutional attention block stacked network (AO-CABNet) is proposed to replace the receiver to improve the ability to recover the original signal. Simulation results show that the AO method has better optimization capability to quickly obtain the optimal parameters of the network model, and the proposed AO-CABNet structure outperforms existing schemes.

key words: *OFDM, convolutional neural network, attention mechanism, Aquila optimizer*

1. Introduction

With its advantages of strong interference immunity and high spectral efficiency, orthogonal frequency division multiplexing (OFDM) has developed into a promising technical solution for achieving high-speed underwater acoustic (UWA) communications [1]. However, OFDM suffers from high peak-to-average power ratio (PAPR) problem [2]. Many studies aim to reduce the PAPR, where clipping [3] is simple to implement but leads to in-band distortion and out-of-band noise, which affects the receiver performance.

Recently, deep learning (DL) has been introduced to wireless communication. In the case of channel distortion, Ye et al. [4] propose a fully connected deep neural network (FC-DNN) to process wireless OFDM channels, implicitly estimating channel state information (CSI) and directly recovering faded signals. For UWA-OFDM systems, the FC-DNN scheme covers the channel decoding, channel estimation, signal detection and constellation demapping modules with improved performance [5]. Considering that FC-DNN has high sensitivity to distortion and is not suitable for complex UWA channels, [6] applies the convolutional neural network (CNN) to the physical layer to extract potential features from the received signal and cascade multilayer perceptron (MLP) demodulation to recover the faded signal. However, it ignores the outliers caused by the UWA environment and pays excessive attention to those invalid data, which not only reduces the efficiency, but also degrades the detection results.

In order to solve the above problems, we propose a network consisting of CNN-attention blocks (CAB) to as a receiver for UWA-OFDM systems, jointly accomplishing channel estimation and signal detection. This network integrates CNN and attention mechanism (AM) [7] at each layer to extract signal sequence features and learn them autonomously, and finally cascades the MLP to complete the recovery of fading signals. Considering the problem of poor network generalization ability during training, Aquila optimizer (AO) [8] is used to obtain the optimal parameters.

Manuscript received July 18, 2023.
Manuscript revised October 24, 2023.
Manuscript publicized December 1, 2023.
[†]College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China.
a) E-mail: liufeng@shmtu.edu.cn
DOI: 10.1587/transfun.2023EAL2065

2. System Model

Inspired by the existing DL model [9], we propose the AO-CABNet based clipped UWA-OFDM communication as shown in Fig. 1. The data stream $\boldsymbol{b}$ is modulated to signal $\boldsymbol{s}$. In order to make the discrete time domain signal better approximate the continuous signal, the signal is L-fold oversampled to obtain the frequency domain signal $X(K)$. Then the signal is converted from the frequency domain to the time domain using the inverse fast Fourier transform (IFFT) with N points as

$$x(n) = IFFT(X(K)) = \frac{1}{\sqrt{LN}} \sum_{K=0}^{LN-1} X(K)\, e^{j\left(\frac{2\pi Kn}{LN}\right)} \tag{1}$$

The PAPR of an OFDM signal is defined as the ratio of the maximum instantaneous signal power to the average signal power and can be expressed as

$$PAPR(x(n)) = 10\log_{10} \frac{\max\left(|x(n)|^2\right)}{E\left(|x(n)|^2\right)} \tag{2}$$

When the amplitude of the OFDM signal exceeds some specified threshold level, the clipping approach is used to reduce the PAPR as

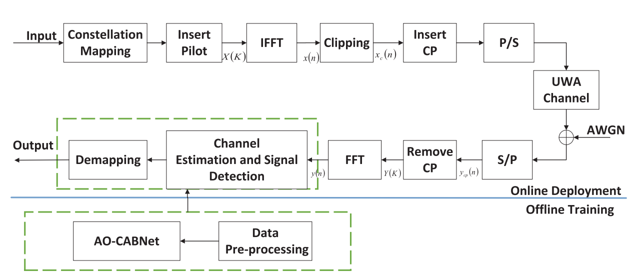

Fig. 1 AO-CABNet based clipped UWA-OFDM system.

$$x_c(n) = \begin{cases} x(n), & \text{if } |x(n)| \le A \\ Ae^{j\Phi_n}, & \text{other} \end{cases} \tag{3}$$

where Φ_n is the phase of $x(n)$ and A is the signal amplitude threshold, expressed as

$$A = CR \cdot \sigma \tag{4}$$

where clipping ratio (CR) is a parameter in the clipping process, and σ is the average power of OFDM symbols. In the traditional clipping method, the value of CR determines the clipping level of the signal peak, which affects the PAPR reduction performance and bit-error-rate (BER) performance.

To eliminate the inter-symbol interference, cyclic prefix (CP) is inserted before each OFDM symbol. After the OFDM modulation, the signal with CP $X_{cp}(n)$ is fed into the UWA channel. The received signal $y_{cp}(n)$ is represented as

$$y_{cp}(n) = x_{cp}(n) \otimes h(n) + u(n) \tag{5}$$

where $\otimes$ denotes the convolution, $u(n)$ denotes the additive white Gaussian noise (AWGN), and $h(n)$ denotes the impulse response of the UWA channel.

At the receiver of the UWA-OFDM system, the channel estimation and signal detection are used to recover the original source data stream $\boldsymbol{b}$.

3. Proposed DL-Based Scheme

3.1 Network Architecture

Considering that AM has strong temporal feature extraction capability and avoids the long-term dependency problem compared to recurrent neural networks (RNN), we integrate CNN and AM to simultaneously learn signal sequence features at multiple levels without affecting each other.

As shown in Fig. 2, the CABNet consists of four layers. In order to accelerate convergence and prevent over-fitting, a layer normalization (LN) layer is added to the network, and the data processed in multiple iterations is finally mapped to the final result by two fully connected layer. The CAB contains two parts: the CNN and the AM as [10]. A typical CNN structure consists of several alternately connected convolutional and pooling layers. To ensure the accuracy of the output signal of the model, the CNN used here includes only the convolutional layer, which does not output the exact value like the hidden layer, but extracts the intrinsic connection between the signal features, i.e., the input of the next layer is the output of the previous layer [11]. Therefore, the output of the n^{th} layer of the CNN can be expressed as

$$\hat{X}_n = f(W_n \otimes Y_n + b_n) \tag{6}$$

where W_n is the weight matrix, b_n is the bias vector, Y_n is the input data, and $f(\cdot)$ denotes the activation function.

Then, AM is integrated after each convolutional layer to extract the time series features by the softmax function as

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}} V\right) \tag{7}$$

where Q, K, V represent query, key, and value respectively, and the d_k is dimension of keys.

The integrated CNN and AM modules process the received signals. After multiple layers of CAB iterations, the CABNet is constructed to learn the global features of the signal sequences. Finally the cascade has an MLP that serves as a classifier to complete the recovery of the decayed signals.

3.2 AO Improved CABNet Algorithm

The structural parameters of the network are crucial for the performance. For instance, the model's generalizability is directly correlated with the number of hidden units [12]. In order to achieve the best training accuracy, different parameters must be tried to construct the CABNet, which include different number of filters, convolutional kernel size and number of hidden units in the fully connected layer. For the CABNet, we adopt the AO to optimize the network structure and obtain the best parameters to be applied to the network, avoiding the tedious and wrong methods.

The AO is a state-of-the-art intelligent optimization algorithm that simulates the behavior of Aquila during the hunting process. The specific steps of the AO improved CABNet algorithm are shown in Algorithm 1. The algorithm involves specific parameters and formula operations as in [8], which are omitted here due to space limitation.

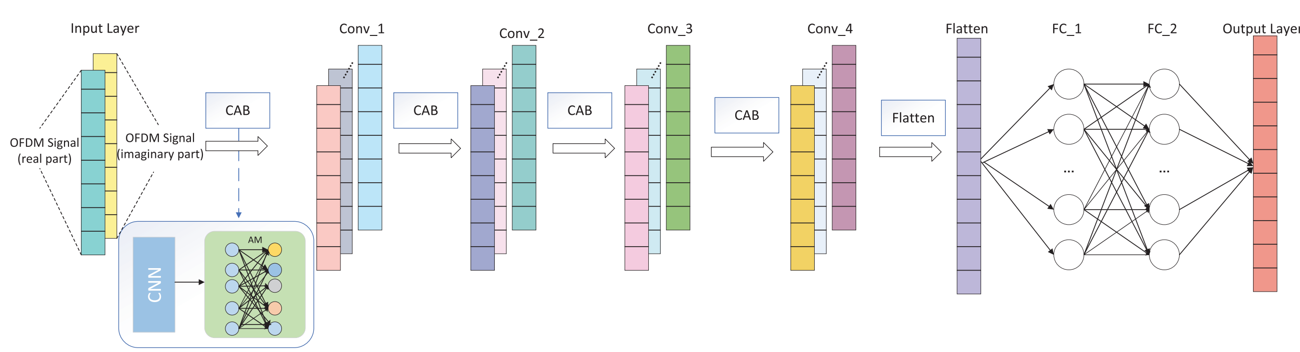

Fig. 2 CABNet structure diagram.

Algorithm 1 Proposed AO improved CABNet algorithm

Initialize parameters, such as N, T, UB_j, LB_j, etc.
Define the fitness function as the prediction accuracy of the network model multiplied by 100.
Initialize the population X and calculate the fitness value.
while $t < T$ **do**
  Find the optimal individual $X_{best}(t)$.
  Calculate the equations for the parameters of each stage.
  for $i = 1, 2, \dots, N$ **do**
    Calculate the mean value $X_M(t)$.
    if $t \le \frac{2}{3} * T$ **then**
      if $rand \le 0.5$ **then**
        $X(t+1) = X_{best}(t) \times \left(1 - \frac{t}{T}\right) + X_M(t) - X_{best}(t) * rand.$
      else
        $X(t+1) = X_{best}(t) \times Levy(D) + X_R(t) + (y - x) * rand.$
      end if
    else
      if $rand \le 0.5$ **then**
        $X(t+1) = \delta \times ((UB - LB) \times rand + LB) - rand + \alpha \times (X_{best}(t) - X_M(t)).$
      else
        $X(t+1) = QF(t) \times X_{best}(t) + rand \times G_1 - (G_1 \times X(t) \times rand) - G_2 \times Levy(D).$
      end if
    end if
    if $Fitness(X(t+1)) > Fitness(X(i))$ **then**
      $X(i) = X(t+1);$
    end if
  end for
end while
return $X_{best}(t), Fitness(X_{best}(t))$

3.3 Training

In the training phase, the data obtained from the simulation of the clipped UWA-OFDM system is firstly pre-processed. In order to improve the signal detection accuracy, the modulated signals are labeled as training data labels, the combined pilot signal and the received signal are used as sample data, and divided into real and imaginary parts combined into two-dimensional data. In this letter, the cross-entropy function is adopted as the loss function, which is shown as:

$$Loss_{CE} = -\sum_{k} y'_k \log(y_k) \tag{8}$$

where y'_k is the output of the model and y_k is the real label. In order to accelerate the convergence speed and improve the training efficiency, the Nadam optimization algorithm is adopted to update the gradient in training and adaptively adjust the learning rate.

4. Simulation Results

4.1 Parameter Setting

During the simulation experiments, we use the BELLHOP ray model [13] to generate the impulse responses of UWA channels. The involved parameters are as shown in Table 1, while the clipped OFDM system parameters are shown in Table 2. Considering the impact of CR on PAPR performance and BER, we chose the clipping with CR of 1.2, 1.4 and 1.6 for the simulation and obtained the CCDF and BER corresponding to the three cases. As shown in Fig. 3 and Fig. 4, the small CR value increases PAPR performance and degrades BER performance. When CR value is large, BER performance degradation can be reduced to a certain extent, but PAPR reduction effect is worse. In order to balance the performance of PAPR and BER, CR of 1.4 is chosen in this paper. The clipped OFDM system is simulated to generate the QPSK and 16QAM datasets required for the experiments. 20,000 simulated data and 100,000 simulated data are obtained from the UWA-OFDM model, respectively, which are divided into the training set and the test set according to the ratio of 4:1.

Table 1 Simulation parameters of the UWA channel.

Simulation parameter	Value
Sending and receiving distance (m)	1000
Sediment layer sound speed (m/s)	2000
density of water (kg/m^3)	1021
density of sediment layer (kg/m^3)	1810
Sea depth (m)	100
Sound source depth (m)	20
Depth of receiving hydrophone (m)	20

Table 2 Simulation parameters of the UWA-OFDM.

Simulation parameter	Value
Number of subcarriers	256
FFT size	1024
Type of modulation	QPSK,16QAM
Oversampling rate	4
CP length	256
Noise model	AWGN

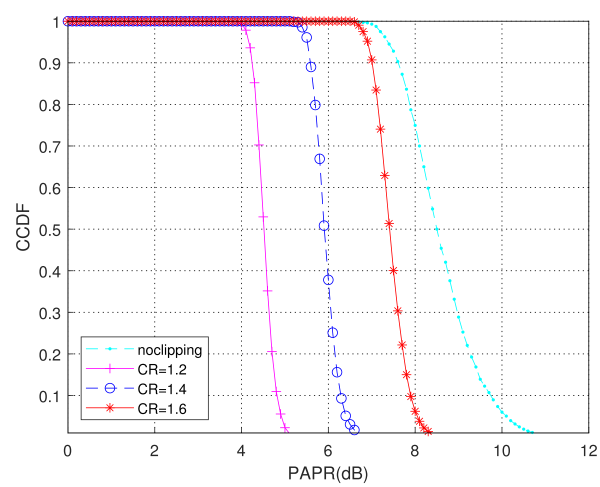

Fig. 3 CCDF comparison of clipping scheme (CR=1.2, 1.4 and 1.6).

4.2 AO Performance

In order to verify the convergence speed and optimization performance of the AO algorithm, the genetic algorithm

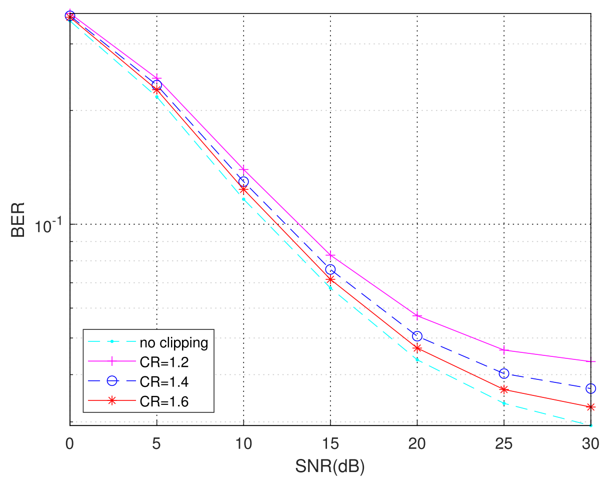

Fig. 4 BER comparison of clipping scheme (CR=1.2, 1.4 and 1.6).

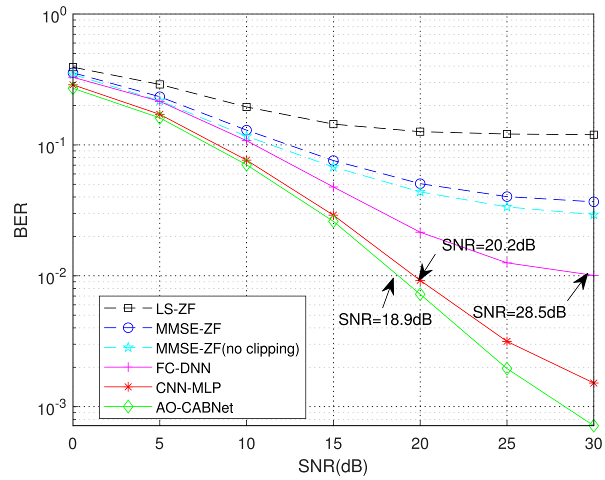

Fig. 6 BER performance comparison with QPSK.

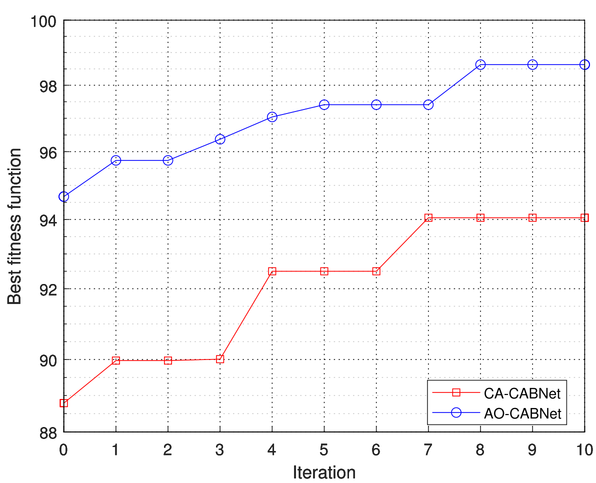

Fig. 5 Comparison of fitness functions of AO and GA.

Table 3 Optimization results.

Algorithm	Number of filters	kernel size	Hidden layer	Fitness
GA	[175,97,175,8]	[2,5,2,3]	[551,495]	94.05
AO	[134,75,134,8]	[2,2,2,3]	[518,490]	98.63

(GA) was adopted as a benchmark to optimize the CABNet under the same environment parameters and dataset. The curves of the fitness function and the optimization results obtained by the two optimization algorithms were compared.

The fitness function value curves are shown in Fig. 5. As can be seen from Fig. 5, the AO optimizes the CABNet parameters with the highest value of the fitness function, faster convergence, and better performance than the GA. This is because four methods of updating population positions are adopted in the AO, which makes the range of exploration more flexible and thus improves the optimization efficiency of the algorithm.

Specific optimization results, the best individuals obtained by the AO and GA and the corresponding highest fitness function values are shown in Table 3.

4.3 BER Performance

BER performance versus signal-to-noise ratios (SNR) should be demonstrated. For comparison, we chose the following methods:

- LS/MMSE-ZF: Conventional least squares (LS), minimum mean square error (MMSE) based channel estimation [14] and Zero Forcing (ZF) signal detection [15] for clipping system.
- MMSE-ZF (no clipping): MMSE-ZF is taken for channel estimation and signal detection for non-clipping system.
- FC-DNN: [4] proposed typical DL-based receiver.
- CNN-MLP: [6] proposed CNN-based receiver.

After data pre-processing, the training set data is first taken for offline training of the network model, and then the completed training model is deployed online to recover the transmitted data stream. The performance results with QPSK and 16QAM are shown in Fig. 6 and Fig. 7, respectively.

As Fig. 6, since QPSK is only phase rotation modulation, the effect of clipping is approximately linear distortion. Compared with BER obtained with MMSE-ZF (no clipping), the performance of MMSE-ZF is slightly reduced. With or without the use of clipping techniques, the DL-based methods are significantly better than the conventional methods. This is because the traditional methods often ignore the influence of underwater acoustic channel noise, resulting in a poor BER. Among the three DL methods, the CNN-MLP performs better compared to FC-DNN due to the strong feature extraction capability of CNN, which ensures accurate signal recovery. However, the performance of AO-CABNet is even better. The performance gain compared to FC-DNN and CNN-MLP is approximately 9.6 dB and 1.3 dB at the BER of 10^{-2}. Therefore, the AO-CABNet not only has stronger signal sequence feature extraction and learning ability, but also can focus on effective feature information, suppress abnormal features, and more effectively compensate

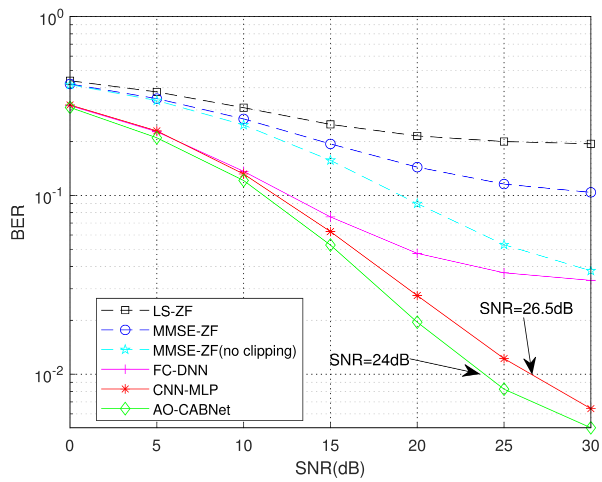

Fig. 7 BER performance comparison with 16QAM.

for nonlinear distortion during signal transmission.

Figure 7 shows similar results. Since 16QAM is amplitude modulated and phase modulated, the effect of clipping can be viewed as nonlinear distortion. The performance of MMSE-ZF is significantly degraded compared to the BER obtained with MMSE-ZF (no clipping). However, the DL-based methods still show superiority over the traditional ones. In high SNR region, the proposed method achieves the best performance with larger SNR gain than the other case. This verifies that the proposed method has strong robustness.

5. Conclusion

In this letter, a DL-based network model was proposed to replace the receiver of a clipped UWA-OFDM system, which can effectively combat the nonlinear noise generated during the clipping process and thus improve the ability of the system to recover the faded signal. The model integrates AM and CNN to form the CABNet to extract signal sequence features, and uses the newly proposed intelligent optimization algorithm AO to optimize the network parameters. Simulation results show that for the optimization process, the AO converges faster and has better optimization capability than the GA. The AO-CABNet outperforms the traditional method and the existing DL methods in the joint channel estimation and signal detection process.

Acknowledgments

This research was funded by the Innovation Program of Shanghai Municipal Education Commission of China (2021-01-07-00-10-E00121), the Tibet Autonomous Region Science and Technology Program (CGZH2024000285), and the Natural Science of China (62271303).

References

[1] B. Li, S. Zhou, M. Stojanovic, L. Freitag, and P. Willett, "Multicarrier communication over underwater acoustic channels with nonuniform Doppler shifts," IEEE J. Ocean. Eng., vol.33, no.2, pp.198–209, 2008.

[2] S.H. Han and J.H. Lee, "An overview of peak-to-average power ratio reduction techniques for multicarrier transmission," IEEE Wireless Commun., vol.12, no.2, pp.56–65, 2005.

[3] T.N. Munni and M. Hossam-E-Haider, "Performance analysis of peak to average power ratio (PAPR) reduction techniques in OFDM system for different modulation schemes," 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (iCEEiCT), pp.605–610, 2018.

[4] H. Ye, G.Y. Li, and B.H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," IEEE Wireless Commun. Lett., vol.7, no.1, pp.114–117, 2018.

[5] J. Zhang, Y. Cao, G. Han, and X. Fu, "Deep neural network-based underwater OFDM receiver," IET Communications, vol.13, no.13, pp.1998–2002, 2019.

[6] Y. Zhang, C. Li, H. Wang, J. Wang, F. Yang, and F. Meriaudeau, "Deep learning aided OFDM receiver for underwater acoustic communications," Applied Acoustics, vol.187, p.108515, 2022.

[7] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, L.u. Kaiser, and I. Polosukhin, "Attention is all you need," Advances in Neural Information Processing Systems, I. Guyon, U.V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, eds., Curran Associates, 2017.

[8] L. Abualigah, D. Yousri, M. Abd Elaziz, A.A. Ewees, M.A. Al-Qaness, and A.H. Gandomi, "Aquila optimizer: a novel meta-heuristic optimization algorithm," Computers & Industrial Engineering, vol.157, p.107250, 2021.

[9] 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.

[10] K. Ren, S. Yuan, C. Zhang, Y. Shi, and Z. Huang, "CANET: A hierarchical CNN-attention model for network intrusion detection," Computer Communications, vol.205, pp.170–181, 2023.

[11] H. Yang, X. Zhang, A. Yi, R. Wang, B. Lin, H. Xing, and B. Sha, "A modified convolutional neural network-based signal demodulation method for direct detection OFDM/OQAM-PON," Optics Communications, vol.489, p.126843, 2021.

[12] Q. Liu, N. Cao, M. Li, Z. Zhang, and G. Yang, "Wireless channel state prediction method based on improved adaptive and parameter-free recurrent neural structure," IEEE Access, vol.10, pp.63329–63338, 2022.

[13] M.B. Porter, "The bellhop manual and user's guide: Preliminary draft," Heat, Light, and Sound Research, La Jolla, CA, USA, Technical Report, vol.260, 2011.

[14] F. Zhong and W. Zhou, "Evaluation of channel estimation algorithms in OFDM underwater acoustic communications," Proc. 10th International Conference on Underwater Networks & Systems, WUWNet'15, New York, NY, USA, Association for Computing Machinery, 2015.

[15] K. Ramadan, M. Dessouky, S. Elagooz, M. Elkordy, and F. Abd El-Samie, "Equalization and carrier frequency offset compensation for underwater acoustic OFDM systems," Ann. Data Sci., vol.5, no.2, pp.259–272, 2018.