

Article

Underwater-Acoustic-OFDM Channel Estimation Based on Deep Learning and Data Augmentation

Jiasheng Guo ^{1,2}, Tieliang Guo ^{2,3,*}, Mingran Li ¹, Thomas Wu ¹ and Hangyu Lin ^{1,2}¹ School of Computer, Electronics and Information, Guangxi University, Nanning 530004, China² Guangxi Key Laboratory of Machine Vision and Intelligent Control, Wuzhou University, Wuzhou 543002, China³ Academy of Sciences, Wuzhou University, Wuzhou 543002, China

* Correspondence: guotieliang@hrbeu.edu.cn

Abstract: In UnderWater-Acoustic-Orthogonal-Frequency-Division-Multiplexing-(UWA-OFDM) communication, the traditional interpolated channel estimation method produces error codes, due to the small number of user pilots, uneven distribution, and complex channel characteristics. In this paper, we propose a novel UWA-channel-estimation method based on Deep Learning (DL). First, based on a small number of channel samples, we used the CWGAN-GP model to generate enhanced classified underwater-acoustic channel samples to have semantic similarity to the real samples and also to present the diversity of the samples. After obtaining the channel sample, the pilot estimation matrix was processed in a similar image way. Here, we extracted the channel features by constructing a convolutional network structure similar to U-Net, weakening the impact of feature information loss. A Channel-Attention-Denoising-(CAD) module was also designed, to further optimize the reconstructed channel information. The simulation results verified the superiority of the proposed algorithm, in terms of Mean Square Error (MSE) and Bit-Error Rate (BER) compared to the existing Least-Squares-(LS), Deep-Neural-Network-(DNN), and ChannelNet algorithms.

Keywords: underwater acoustic; OFDM; DL; channel estimation; channel attention



Citation: Guo, J.; Guo, T.; Li, M.; Wu, T.; Lin, H. Underwater-Acoustic-OFDM Channel Estimation Based on Deep Learning and Data Augmentation. *Electronics* **2024**, *13*, 689. <https://doi.org/10.3390/electronics13040689>

Academic Editor: Christos J. Bouras

Received: 31 December 2023

Revised: 30 January 2024

Accepted: 1 February 2024

Published: 7 February 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Among the currently known forms of energy radiation, the sound wave is the best carrier of underwater wireless communication [1,2]. The radio communication mode usually used on land is that of electromagnetic waves as the carrier. When the electromagnetic waves propagate in the water, they are absorbed in large quantities and are rapidly attenuated, and the action distance is very limited. Sound waves have good propagation performance underwater, and the attenuation coefficient of sound waves ranging between 1 Hz and 50 kHz in water ranges from about 10^{-4} dB/m to 10^{-2} dB/m [3]. After years of development, Underwater-Acoustic-(UWA) communication technology has undergone a transition from non-coherent communication to coherent communication. Additionally, multiple-access communication methods, such as Orthogonal Frequency-Division Multiplexing (OFDM), represent a shift from single-carrier communication. These advancements reflect the continuous pursuit of higher communication rates and bandwidth requirements in underwater communication [4–6]. The complex and changeable characteristics of UWA transmission lead to a series of problems in the UWA channel—such as strong multiple paths, fast decline, and complex noise interference—which present great difficulties and challenges for underwater information transmission [7,8]. Estimation and equalization can effectively obtain the channel information and restore the transmitted signal at the receiving end. This is an important part of underwater comprehensive information perception and information interaction, and has very important research significance.

In a traditional UWA communication system, it is usually necessary to transmit a fixed pilot sequence at the transmitting end, and to accurately estimate the channel at the

receiving end, so as to realize the recovery of the data signal. Common channel-estimation and equalization methods include the Least-Squares-(LS) algorithm [9,10], the Minimum-Mean-Square-Error-(MMSE) [11] algorithm, and the Compressed-Sensing-(CS) method [12]. Although the LS algorithm is simple to implement, its overall performance is poor. While the MMSE algorithm can approximate the optimal solution in a statistical sense, it requires a prior channel state as the premise, so its operability in a practical application is not strong. In addition, this kind of algorithm does not have the characteristics of adaptation. The CS algorithm can utilize the sparsity of the UWA channel to reconstruct the original signal at a sampling rate lower than the Nyquist sampling rate. Keeping this in consideration, the authors in [12] formulated an iterative-UWA-channel-estimation approach, assuming that the UWA channel undergoes Rayleigh fading.

In recent years, the design of a UWA communication system based on machine learning has been an emerging research topic, and machine-learning methods have great potential to promote the further development of UWA communication technology. In [13], the authors established a robust five-layer-neural-network model designed for water-acoustic channel estimation and equilibrium. They applied this model to simulate the channel, using BELLHOP for offline training and online testing. The outcomes revealed that the Deep-Neural-Network-(DNN)-based network model exhibited significant advantages over traditional algorithms, particularly in scenarios with limited OFDM-communication pilot numbers. In [14], the authors also proposed a joint UWA-OFDM demodulation, channel-estimation, and balanced-neural-network structure, to realize the integration of the receiving device. The authors in [15] proposed a UWA receiving system based on a Deep Belief Network (DBN), which effectively alleviated the performance decline caused by the Doppler effect and multiple-path transmission in communication. The simulation and offshore test results verified the effectiveness of this method. Unlike traditional methods, two different framework models based on DNN were proposed in [16], to address the UWA-OFDM-channel-estimation problem. Extensive experiments were conducted on the proposed methods, to compare them to traditional methods, such as LS, MMSE, and the back propagation algorithm. Although previous research has demonstrated the effectiveness of machine-learning methods in improving the overall performance of UWA channel estimation, there are still many key problems to be solved, regarding the specific application scenarios of UWA. One major challenge is the scarcity of samples. UWA communication faces limitations imposed by sea test conditions and other factors, resulting in low collection efficiency in the UWA environment and high acquisition costs for communication data. Therefore, with the sample size collected by UWA communication in a limited time, it is usually difficult to support the effective training of the model, thus causing overfitting of the model. Another hurdle is data-calibration difficulties. Existing models usually use the real channel impulse response as the training label. However, it is difficult to obtain the true value of the channel in real UWA communication in real time—that is, the calibration of the training labels cannot be realized, so the learning of the model usually needs to be conducted offline. Hence, there is a pressing need for research focused on enhancing UWA channel data.

Traditional data-enhancement algorithms include replay technology based on the statistical characteristics of raw water-acoustic-channel-distribution and data-enhancement methods based on communication-signal-processing technology. Channel replay is able to generate data with the same statistical properties, based on a small number of measured UWA channels. Relevant studies have proposed a UWA channel simulator driven by measurement data, which assumes that in a specific time window the communication channel is a stochastic process [17]. In [18], the authors used Empirical-Mode-Decomposition-(EMD) technology to first dismantle the channel into the true part and the random part [19]. On the basis of ensuring the inherent characteristics of the UWA channel, the method extends the latter and recombines with the confirmed part, to generate new UWA channel data. With the help of disturbances and interference often occurring in UWA communication scenarios, such as synchronization error, Doppler offset, noise interference, etc., data enhancement

based on communication-signal-processing technology can realize data expansion. Typical work includes applying a symbolic time shift to the raw data in the literature and a Doppler shift to achieve data enhancement [20].

Data-enhancement algorithms are needed in many application fields of deep learning, and the research from the perspective of deep learning provides a way for the reference of UWA channel estimation. Typical of these is Generative-Adversarial-Networks-(GANs) data enhancement. GANs have the powerful ability to produce simulated samples that capture the actual data distribution [21]. The training algorithm of GANs for generating simulated samples highly similar to the actual data distribution essentially involves a two-player minimax game with two adversarial networks and one discriminative network. The generator network G uses random latent variables, following a standard Gaussian distribution or a uniform distribution, to generate simulated samples with the goal of deceiving the discriminator network. Conversely, the discriminator network D aims to differentiate between the simulated samples generated by the network G and real samples that adhere to the actual data distribution. This implies that a trained GAN can effectively capture the complex distribution of a UWA channel. By using a GAN's generator network, we can obtain a large number of generated samples of simulated UWA channels. However, the training of GANs is usually unstable, with the problems of network convergence, mode collapse, and gradient vanishing [22], etc. Therefore, in this paper we propose a novel GAN variant to improve the training stability. After obtaining a rich sample of UWA channels, channel estimation is performed on the received signals passing through the UWA channels, using our proposed DL framework.

In this paper, we propose a DL architecture to implement channel-sample generation and channel estimation for UWA-OFDM communication. The main contributions of this study are summarized as follows:

- (1) we propose a DL-based UWA-OFDM-channel-estimation scheme for the defects in the traditional UWA-OFDM communication system. We consider the shortage of underwater-acoustic channel samples and we introduce the CWGAN-GP model to generate enhanced channel samples, whose generator contains some channel feature parameters to realize different types of channel model generation.
- (2) We propose a novel deep-learning channel-estimation architecture. First, in order to attenuate the effect of feature information loss, we extracted channel features by constructing a convolutional network structure in the way of U-Net, and we then designed a Channel-Attention-Denoising-(CAD) module, to further optimize the reconstructed channel information.
- (3) We provide a validation of the channel quality in different UWA environments. The numerical results show that the proposed model is better than the traditional OFDM communication system, DNN-based, and ChannelNet-based models, in terms of Mean Square Error (MSE) and Bit-Error Rate (BER). Significant performance improvements are particularly evident in harsh UWA environments.

The rest of this paper is organized as follows: In Section 2, we briefly introduce the UWA-OFDM communication system. In Section 3, we present the proposed deep-learning-based-channel-estimation scheme. In Section 4, we experiment and evaluate the estimated performance. The paper concludes with Section 5.

2. The UWA-OFDM Communication System

We consider a downlink UWA-OFDM communication system, and a schematic representation of the traditional system model is shown in Figure 1. The information $X(k)$ to be transmitted passes through the baseband signal processing to obtain the transmitted signal $x(t)$. Then, $x(t)$ passes through the UWA channel to the receiving device, to obtain $y(t)$. After a series of reverse transformations, the frequency-domain information $Y(k)$ can be obtained. For UWA channels, the general expression is as follows:

$$h(t, \tau) = \sum_{l=1}^L A_l(t) \delta(\tau - \tau_l(t)). \quad (1)$$

Here, we make the assumption that the UWA channel exhibits linear invariance within an OFDM transmission signal. This allows for a satisfactory approximation of the UWA channel as the channel impulse response of the L principal discrete paths. $A_l(t)$ is the decay coefficient of the constant path l , and $\tau_l(t)$ is the time delay corresponding to the path l .

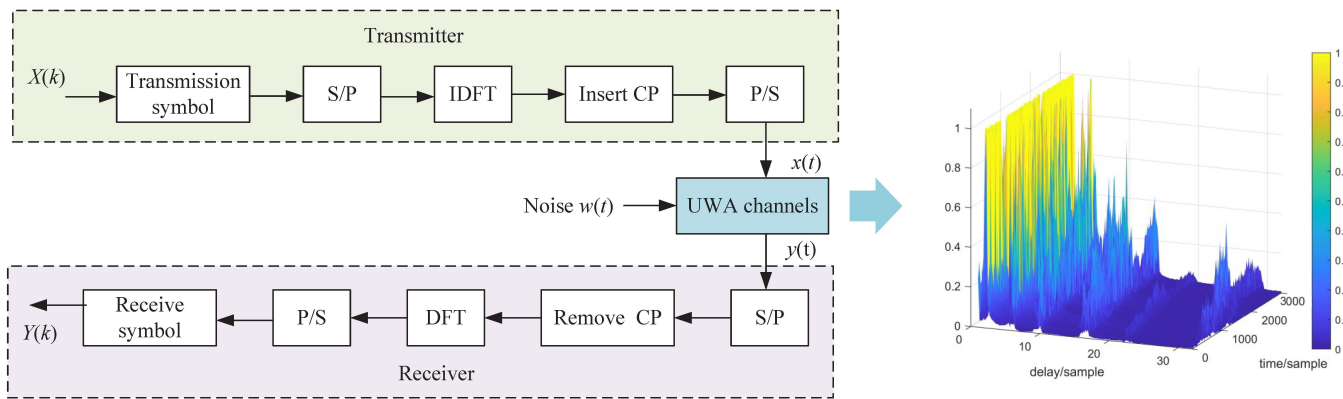


Figure 1. Traditional UWA-OFDM communication system.

The propagation of sound waves in the ocean is affected by the wave fluctuations of the sea surface, the uneven stratification of the seabed, and the scattering and refraction effects caused by the heterogeneity of the seawater medium. In addition, the complexity of a UWA channel is reflected because it changes over time and through space, and these changes occur randomly, so an additional statistical-modeling stage is required. However, some efforts are still being made to build standard models to succinctly describe the statistics of UWA channels [23]. Therefore, in this paper, we used the powerful fitting ability of the DL method to reconstruct the UWA channel of the data block, based on the channel estimated by the pilot block.

At the transmitter, the transmission symbol of the inserted pilot undergoes initial conversion into parallel data streams. Subsequently, it is modulated onto various subcarriers through a Discrete-Fourier-Reverse-Transform-(IDFT) unit. Thereafter, a Cyclic Prefix (CP) is inserted, to mitigate inter-symbol interference. The transmission signal passes through parallel-to-serial conversion and then is sent to the water-acoustic channel. The signal received at the receiver is

$$y(t) = \sum_{l=1}^L A_l(t) x(t - \tau_l(t)) + w(t), \quad (2)$$

where $w(t)$ represents the Additive-White-Gaussian Noise (AWGN). After removing the CP and performing the DFT, the received frequency-domain signal becomes

$$Y(k) = X(k)H(k) + W(k). \quad (3)$$

In the conventional UWA-OFDM system, the pilot signal is extracted and utilized for channel estimation. The LS estimate of $H(k)$ can be represented as [9]

$$\hat{H}(k) = \frac{Y_{pilot}(k)}{X_{pilot}(k)} = H(k) + \frac{W(k)}{X_{pilot}(k)}. \quad (4)$$

Here, we assume that the pilot is distributed in a comb-like structure, to better estimate the effect of the Doppler.

3. The UWA-OFDM-Channel-Estimation Framework Based on Deep Learning

This section outlines the UWA-OFDM-channel-estimation framework based on deep learning. According to the shortcomings of insufficient channel samples, we first used the data-enhancement method to increase the number of channel samples, and we then used deep learning for channel estimation. The overall framework is shown in in Figure 2.

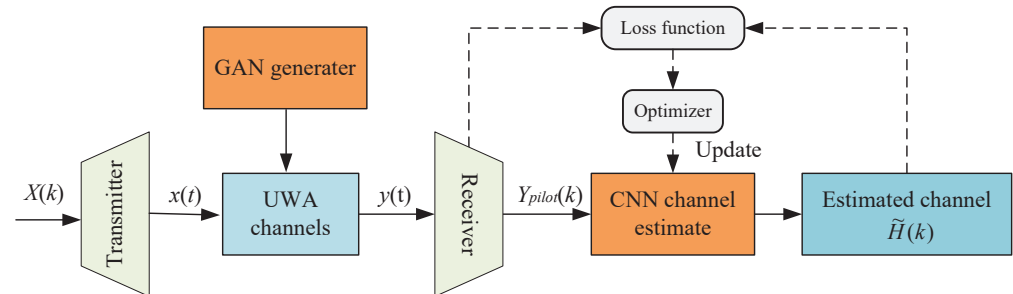


Figure 2. The UWA-OFDM-channel-estimation framework based on deep learning.

3.1. GAN-Based UWA-OFDM Channel Sample Enhancement

The GAN structure is shown in Figure 3, including the generator and the discriminator. In the generator, to learn the distribution p_g on the UWA channel data h , we randomly define a prior noise variable z as input that satisfies a normal distribution or other arbitrary distribution. The generator network $G(z; \theta_g)$ maps the latent distribution p_z of the latent variables z into the data distribution p_g , thus obtaining a simulated sample that fits the distribution p_g , which is a differentiable function characterized by the multilayer perceptron G . In the discriminator, we need to define a multilayer perceptron $D(z; \theta_d)$. $D(z; \theta_d)$ denotes the probability of converting h into a sample of actual data that fits the distribution p_{data} while different from p_g . The role of the generator is to generate false data similar to the real data, while the discriminator is responsible for classifying the real data and the false data. We train D to maximize the probability of correctly assigning true samples and generating samples, so we can train G simultaneously by minimizing the $\log(1 - D(G(z)))$. That is to say, the discriminator D and the generator G play a minimax game on the value function $V(G, D)$:

$$\min_{\theta_g} \max_{\theta_d} V(D, G) = E_{h \sim p_{data}(h)} [\log D(h; \theta_d)] + E_{z \sim p_z(z)} [\log(1 - D(G(z; \theta_g); \theta_d))], \tag{5}$$

where θ_g and θ_d are the training parameters of the network. Furthermore, the generative and discriminant networks alternatively optimize the parameters θ_g and θ_d , based on their respective loss functions. That is, the parameters of one network (i.e., G or D) are calculated by using the error-Back-Propagation-(BP) algorithm and the optimization method (e.g., the gradient-descent method), while the parameters of the other network are fixed. By iteratively optimizing the parameters of the GAN network, the generative and discriminant networks learn reasonable mapping functions. This means that the distribution p_g of the simulated samples generated by the generator network converges to the actual data distribution p_{data} , and the discriminator network cannot distinguish the simulated samples that match the distribution p_g from the actual sample (matching the distribution p_{data}).

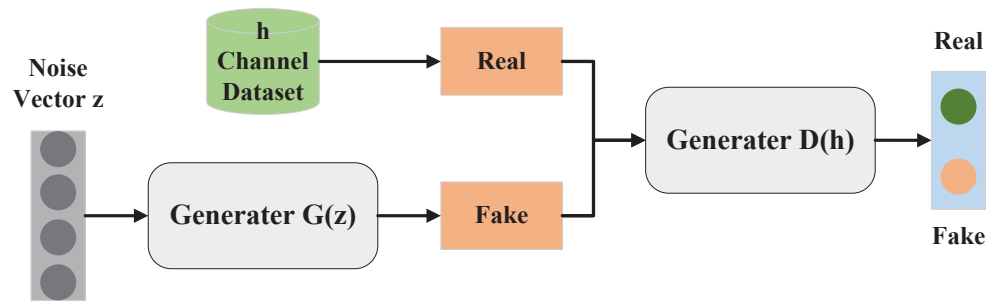


Figure 3. The GAN structure.

GANs are very clever in theory. We should first train the discriminator as well as possible during the training process. But, in practice, the better the discriminator is trained, the more difficult is the generator to optimize, which eventually leads to a GAN training crash. Therefore, we adopted the Wasserstein-GAN-with-Gradient-Penalty-(WGAN-GP) architecture, to alleviate the instability during GAN training [24], as shown in Figure 4. By introducing the Wasserstein distance, which has smoothing properties superior to the Jensen–Shannon (JS) divergence, the gradient-vanishing problem can be theoretically solved. Then, we used a discriminator neural network with a limited numerical range to maximize the mathematical transformation of the Wasserstein distance in a solvable form, and the Wasserstein distance could be approximated. In the regime of this approximate optimal discriminator, optimizing the generator serves to decrease the Wasserstein distance, effectively bringing the generated distribution closer to the real distribution. The WGAN not only addresses the issue of unstable training but also provides a reliable metric for tracking training progress, with this metric being highly correlated to the quality of the generated samples. The WGAN-GP incorporates a gradient penalty [25], replacing the weight-clipping method used in the WGAN. The gradients of the discriminator are restricted by directly imposing an additional penalty term. Therefore, the loss function can be expressed as

$$\min_{\theta_g} \max_{\theta_d} V(D, G) = \mathbb{E}_{h \sim p_{data}(h)} [D(h; \theta_d)] + \mathbb{E}_{z \sim p_z(z)} [-D(G(z; \theta_g); \theta_d)] - \lambda GP(D), \tag{6}$$

where λ is the penalty coefficient, and the last term $GP(D)$ can be written as

$$GP(D) = \mathbb{E}_{\hat{h} \sim p_{sample}(\hat{h})} [(\|\nabla_{\hat{h}} D(\hat{h}; \theta_d)\|_2 - 1)^2], \tag{7}$$

where the distribution $p_{sample}(\hat{x})$ fits the true data distribution p_{data} and the generated data distribution p_g point pairs are evenly sampled along the line. However, for the unconditional generation model, the generation of the data is uncontrollable. We hoped to use the posterior information in the training data, by attaching some channel information of the real data to the model, which could guide the model to generate our desired channel data. The specific generator and discriminator network structures are shown in Figure 5. Eventually, the loss function was rewritten as

$$\min_{\theta_g} \max_{\theta_d} V(D, G) = \mathbb{E}_{h \sim p_{data}(h)} [D(h|c; \theta_d)] + \mathbb{E}_{z \sim p_z(z)} [-D(G(z|c; \theta_g); \theta_d)] - \lambda GP(D). \tag{8}$$

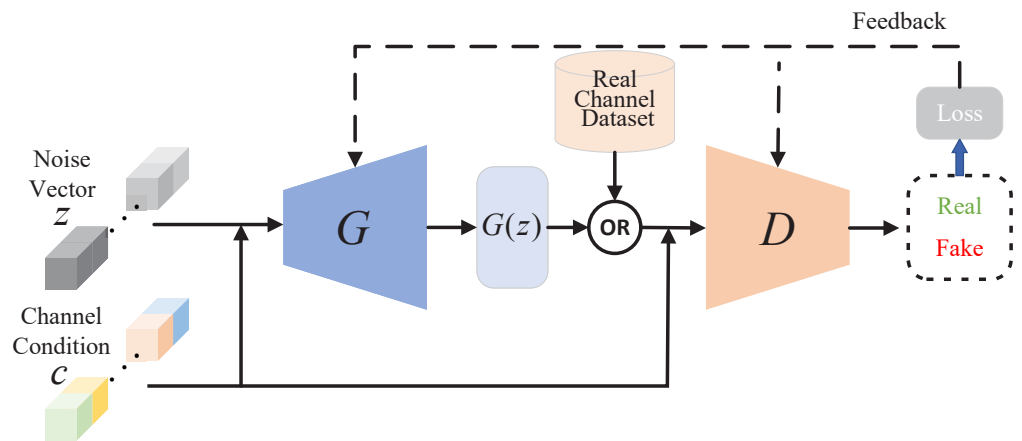


Figure 4. The CWGAN-CP structure.

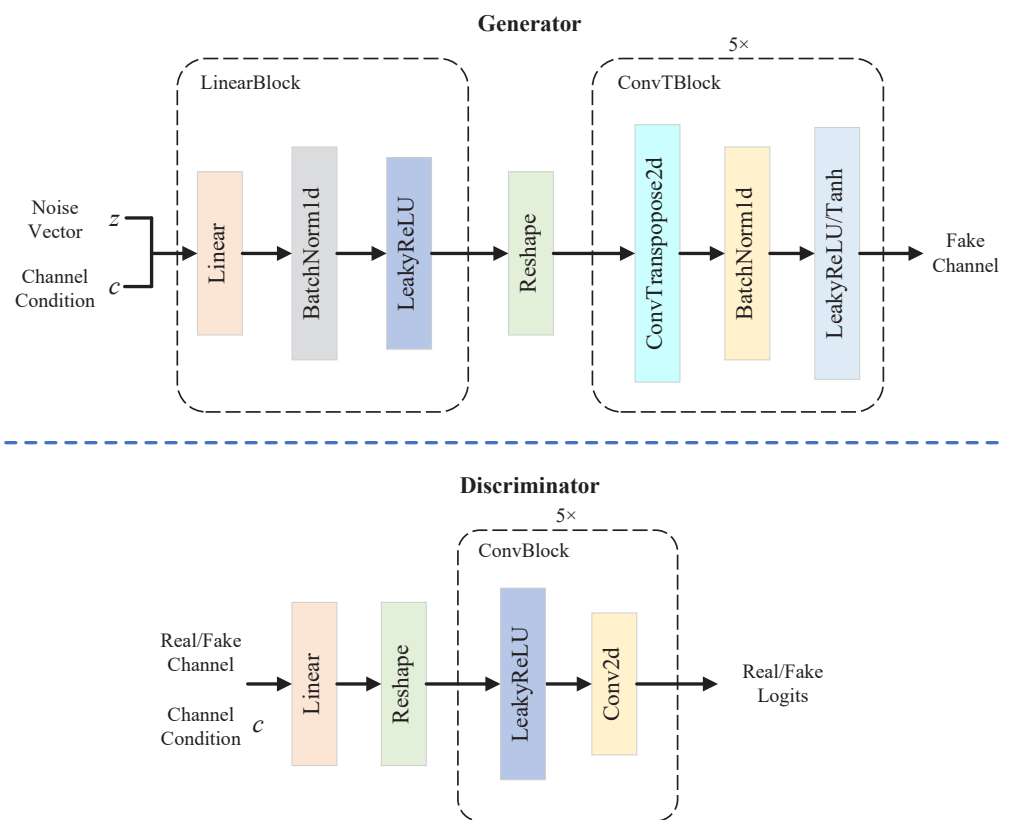


Figure 5. The generator and discriminator network structures.

3.2. DL-Based-Estimated-UWA-OFDM Channel

We used the CNN-based-network framework to estimate the channel state, before which we needed to solve some necessary problems. First, as an input to the network, the image data are usually represented by the RGB three-channel matrix, but the channel-state information is all complex. Therefore, we wanted to modify the input layer of the network for this case. We first divided each complex matrix into two part-matrices: real and imaginary. To preserve the potential correlation between the real and imaginary parts, we reconstructed the two matrices overlapping as a new two-channel matrix. Moreover, some elements in the channel-state-information matrix usually have negative values, which affects the back propagation in the convolutional network and the correct convergence of

the network. To do this, we used a Leak-Rectified Linear unit (LReLU) as our activation function after each convolution layer [26]:

$$\text{LReLU}(x) = \begin{cases} x, & x > 0 \\ b_i x, & x \leq 0 \end{cases} \quad (9)$$

where b_i was the constant parameter between 0 and 1. As the value in the matrix was very close to 0, it was likely to cause gradient vanishing throughout the convolution calculation. In order to solve this problem, we initialized the data before sending them to the network, and we eliminated the influence on the estimated-channel-state-information matrix in the output layer.

The proposed channel-estimation-network framework based on the CNN is shown in Figure 6. First, the network input the data $\hat{H}_{pilot}(k)$ for the preprocessing operation. The input $\hat{H}_{pilot}(k)$ had to be divided into a real part and a virtual part. Then, we used several convolutional layers acting on the process of upsampling and downsampling, respectively, which were used as features to extract the signal and the transmission of internal information. We used the residual structure between the layers [27], and the output of all the previous layers was the input of the next layer. The output of all the previous layers was the input of the next layer. The output $\tilde{H}(k)$ can be expressed as

$$\tilde{H}(k) = F_n(F_{n-1}(\dots F_1(\hat{H}_{pilot}(k)))) \quad (10)$$

where $F_n(\cdot)$ represents convolution operation. Due to constant underground sampling, the feature information eventually passes through the bottleneck layer, where a lot of detail features may be lost. To alleviate this problem, we adopted a U-Net-like structure, and between upper and downsampling, we used the operation of jump connections, which integrated richer higher convolutional feature layers and lower convolutional features, to further compensate for the lost information of downsampling during the encoding stage. Each of the networks contained three upsampling and downsampling operations. In the first downsampling, 32 filters were used, and for each time of the subsequent downsampling, the number of filters was twice as many as before. The output from each layer in the upsampling remained consistent with the downsampling.

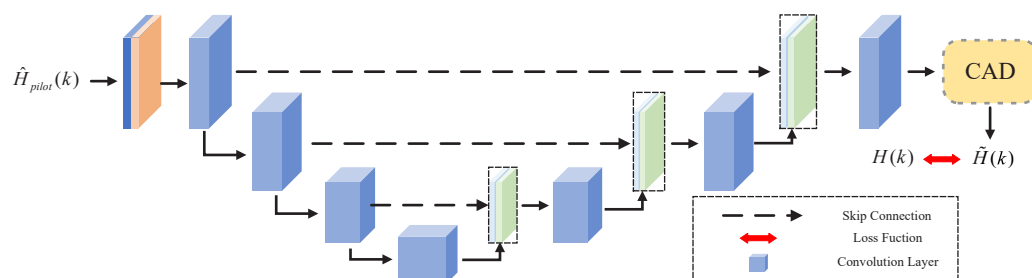


Figure 6. A CNN-based-channel-estimation-network framework.

After obtaining the features oversampled in the last layer, we did not directly regress the channel information, but we designed a Channel-Attention-Denoising-(CAD) module to further optimize the channel information [28], as shown in Figure 7. Here, our module used the channel-attention mechanism, and we wanted the model to selectively focus on or reinforce the information of a specific underwater-acoustic channel when processing features, reducing the attention to unnecessary information. While reducing the model computational burden, the model was allowed to adaptively learn the weights of frequency features during training, as different environments and underwater-acoustic channels may lead to differences in the signal-frequency spectrum. Specifically, we input the features as follows: $f \in \mathbb{R}^{C \times H \times W}$. Then, for each channel i , we calculated the average value:

$$z_i = \frac{1}{H \times W} \sum_{j=1}^H \sum_{k=1}^W f_{i,j,k} \quad (11)$$

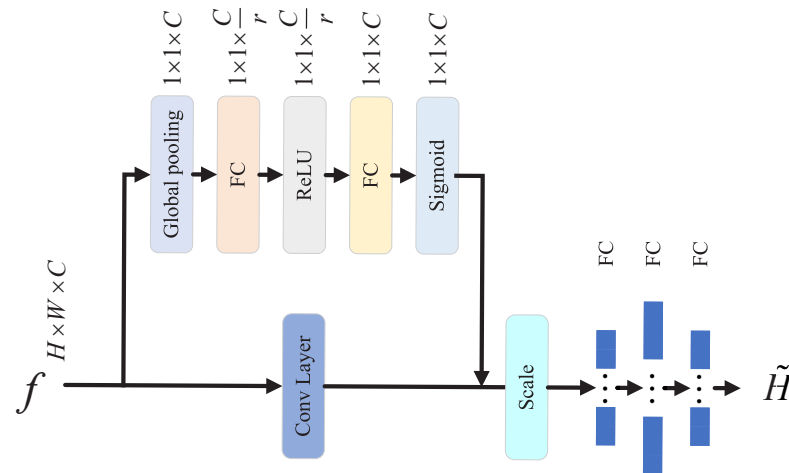


Figure 7. The-channel-attention-denoising module.

Next, we introduced two learnable weight parameters, W_1 and W_2 , mapped into the two representations $u_i = \sigma(W_1 \cdot z_i)$ and $v_i = \sigma(W_2 \cdot z_i)$, where σ was the ReLU activation function. The weighted channel representation $\tilde{f}_i = u_i \cdot X_i$ was obtained by using u and v weighted for each channel, where \tilde{f}_i was a weighted representation of the channel. To ensure that the weighted channel representation and the original input had the same scale, the \tilde{f}_i were normalized as

$$\tilde{f}_i = \frac{u_i \cdot f_i}{\sum_{j=1}^C u_j \cdot f_j} \tag{12}$$

This ensured that the weighted channel representation summed to 1 after normalization. Finally, we multiplied the normalized weighted channel representation with the original input feature map, to obtain the final output:

$$\tilde{H} = \tilde{f} \odot f, \tag{13}$$

where \odot represented the multiplication by element.

The last layer of CNN output comprised data close to the expected value and evaluated by the loss function. Finally, we used MSE and BER as the total loss function of the model, where BER prevented gradient decline when the MSE was small. They were, respectively, denoted as follows:

$$\mathcal{L}_{MSE} = \frac{1}{N} \sum_n (\tilde{H}_n - H_n)^2, \tag{14}$$

and

$$\mathcal{L}_{BER} = \frac{1}{N} \sum_n (\kappa(\hat{Y}_n) - X_n)^2, \tag{15}$$

where \tilde{H}_n and H_n represented the predictive channel information of the model and the corresponding ground-truth values, respectively, \hat{Y}_n and Y_n represented the estimated received information of the model and the corresponding ground-truth values, respectively, and $\kappa(\cdot)$ represented the operation of hard decision making. Therefore, the total loss function was

$$\mathcal{L}_{total} = \mathcal{L}_{MSE} + \mathcal{L}_{BER}. \tag{16}$$

4. Experiments and Evaluations

In this section, we commence by presenting the parameter configurations for the UWA channel and the UWA-OFDM system. We conducted simulation experiments, to assess the authenticity of the channel estimates, the MSE, and the BER performance of our proposed algorithm within the OFDM system.

4.1. Experiments Parameters

In our study, we employed a widely utilized UWA channel simulator developed by [29]. This simulator had undergone validation using real data collected from four experiments, ensuring its ability to generate accurate UWA channels:

$$H(f, t) = \bar{H}_0 \sum_l h_l \bar{\gamma}_l(f, t) e^{-j2\pi f \tau_l}, \quad (17)$$

where \bar{H}_0 was the nominal frequency response, and where h_l and $\bar{\gamma}_l(f, t)$ were the large-scale path gain and the small-scale fading of the l -th UWA path, respectively. Here, we adopted the channels of KAU1 and NOF1 as the simulation environment, and some parameters, summarized in Table 1, to configure the UWA physical environment. Table 2 presents the main simulation parameters of the UWA-OFDM system. For channel samples, the CWGAN-GP generator was used to generate the KAU1 and NOF1 channels, and a CSI matrix of 10,000 pilot positions was obtained after frequency-domain conversion and pilot channel estimation at the receiver. Here, we generated CSI matrices unified at a Signal-to-Noise Ratio (SNR) of 15 dB. Of these, 80% were randomly divided into the training set, with 10% the validation set and the last 10% the test set. Furthermore, we set the scale factor to 10, set the initial learning rate to 0.001, and then decay by 0.1 times every 40 epochs. The maximum epoch was set to 100, but if the value of the loss function did not drop within 5 consecutive epochs, the training was stopped prematurely.

As depicted in Table 2, there were 512 subcarriers, and each frame was composed of 14 symbols, resulting in a CSI matrix size of 512×14 . In our simulation, we evaluated the effectiveness of our proposed algorithm, using four different algorithms: the LS algorithm [10], a DNN-based algorithm [16], the ChannelNet algorithm [30], and FullCSI. The specifics are outlined below:

- *LS*: Select two or four symbols as pilot symbols, and then estimate the CSI of the pilot location and data position by using the LS algorithm and the spline interpolation method, respectively.
- *DNN*: Select two or four symbols as the pilot symbols, and then obtain the CSI matrix by using the DNN-based algorithm.
- *ChannelNet*: Select two or four symbols as the pilot symbols, and then obtain the CSI matrix based on ChannelNet. The algorithm is mainly divided into the process of expansion and denoising reduction.
- *FullCSI*: Suppose that the channel is determined to be known.

Table 1. UWA Channel Parameters.

Parameter	KAU1	NOF1
Water depth	100 m	10 m
Transmitter depth	20 m	10 m
Receiver depth	50 m	10 m
Transmission distance	1080 m	750 m
Spreading factor	1.7	1.7
Sound speed in water	1500 m/s	1500 m/s
Sound speed in bottom	1200 m/s	1200 m/s
Tx vehicle speed	$N(0, 1)$ m/s	$N(0, 1)$ m/s
Rx vehicle speed	0 m/s	0 m/s

Table 2. UWA-OFDM System Parameters.

Parameter	Value
Carrier frequency	8 kHz
Channel bandwidth	6–10 kHz
Number of subcarriers	512
Number of symbols/frame	14
Subcarrier spacing	7.8125 Hz
Symbol duration	128 ms
Cycle prefix length	30 ms
Modulation	QPSK

4.2. Compare the Authenticity of the Channel-Response Estimates

We first verified the fitting ability of the proposed algorithm to the real channel sample after channel estimation. We compared the amplitude of their channel responses under the KAU1 and NOF1 channels separately, as shown in Figure 8. We can see that compared to the NOF1 channel, the KAU1 channel was more complex and the multipath structure was more complex. It was observed that the channel responses estimated by our proposed algorithm could fit well to the real channel samples.

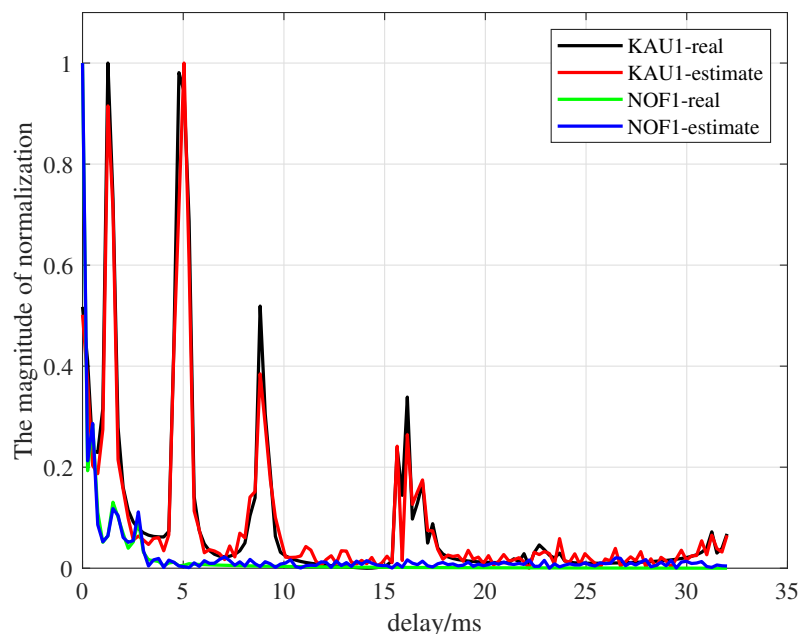


Figure 8. Comparison of the estimated and true channel samples' normalized amplitude in KAU 1 and NOF 1.

4.3. MSE Performance in an Extended Range of SNR

We evaluated the MSE performance over an extended SNR range compared to the training dataset, to investigate the ability to generalize to uncertain SNR values. As shown in Figures 9 and 10, we compared the MSE of several algorithms to different SNR ratios in the KAU1- and NOF1-channel environments. The analysis and simulation results verified that the LS method performs poorly in the case of UWA, because it is inversely proportional to the SNR. While it is easy to implement, this simplicity comes at the cost of relatively low precision. The DNN model was more accurate than the traditional LS algorithm, and its MSE performance at SNR = 30 dB in KAU1 and NOF1 improved by about 184.9%

and 246% compared to LS, because of its strong fitting ability; however, there are also non-convex-optimization and gradient-vanishing problems in fully connected structures, which makes it less stable when handling complex and variable scenarios, such as the simulation-environment NCS1 channel and KAU1 channel. By contrast, the CNN model in the ChannelNet model has strong feature-extraction ability and denoising network, which ensures more accurate channel estimation at lower SNR = -10 dB, such as KAU1 and NOF1, which improved MSE performance by about 175% and 171.4% compared to DNN. Compared to the Channelnet-dimension-expansion network, our proposed algorithm has a stronger feature-extraction ability, and its MSE performance has also been further improved. For real-world applications, it is important that DL-based models have good generalization capabilities, so that they can work efficiently when the online UWA environment does not exactly match the UWA channel used in the training phase. The experimental results show that the NOF1 channel has a simple multipath structure, slow time-varying dynamics, and little improvement. However, in the experiments with KAU1, the model could provide significant performance benefits, due to its excellent channel-reconstruction ability, due to the more challenging channel structure and fast time-varying properties.

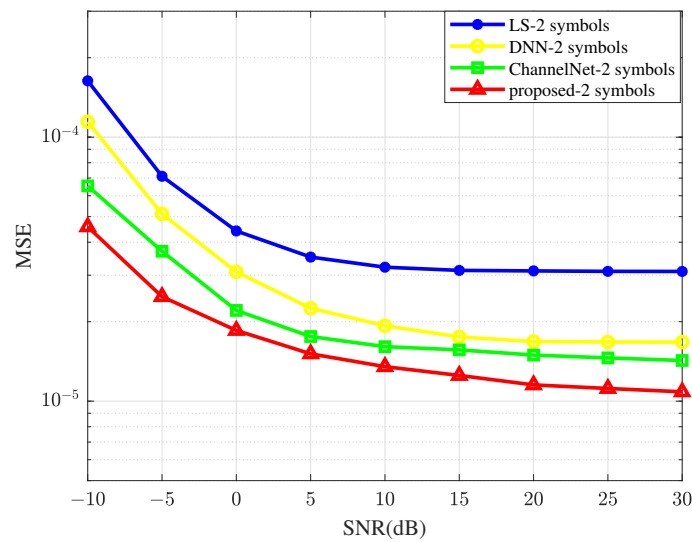


Figure 9. MSE vs. SNR for KAU1 channel.

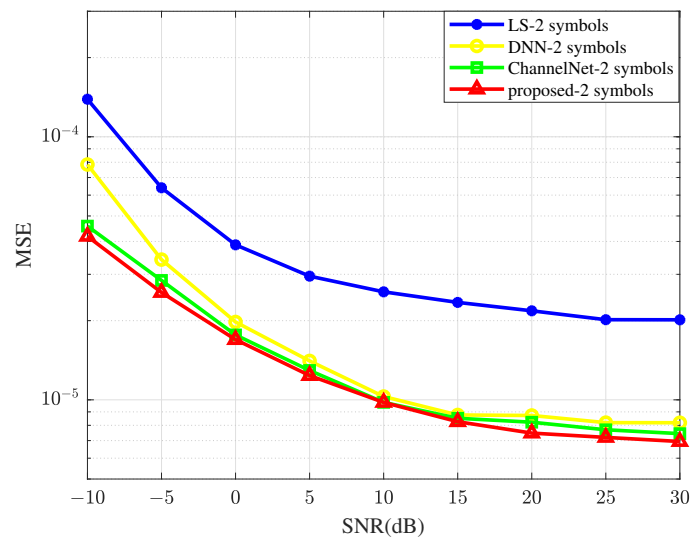


Figure 10. MSE vs. SNR for NOF1 channel.

4.4. BER Performance in an Extended Range of SNR

We also evaluated the BER performance of each method over an extended range of SNR, which corresponded to the data-recovery capabilities. Figures 11 and 12 compare the BER performance for each method, where our proposed method achieved the best performance. The performance of BER depends on the accuracy of the channel-matrix prediction on all data symbols. We can see from the figure that the BER of the proposed algorithm was always lower than that of LS, DNN, and ChannelNet. For example, compared to LS, it improved by about 555.9% at SNR = 30 dB and by about 498.8% at SNR = 20 dB. Similarly, compared to the KAU1 environment, the gap of several deep-learning methods for the NOF1 environment was smaller, and the KAU1 gap for more complex features more obvious.

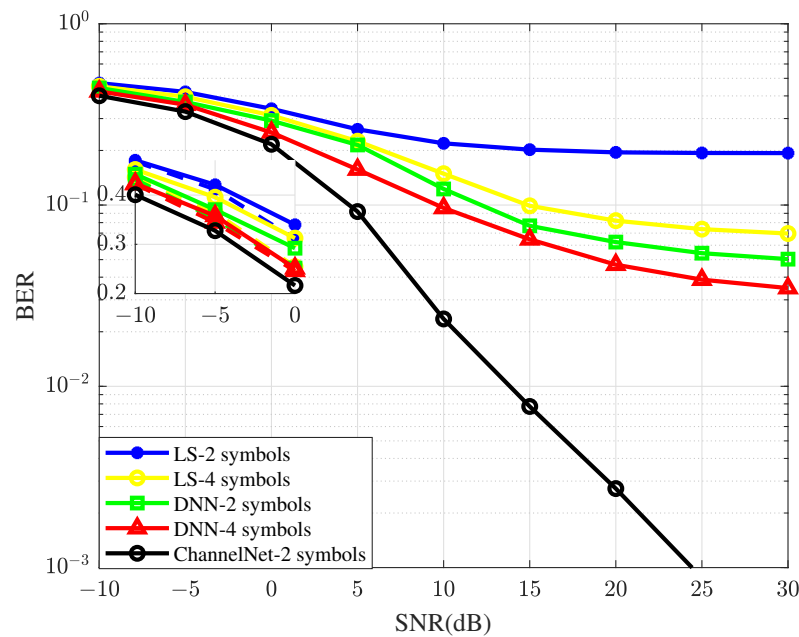


Figure 11. BER vs. SNR for KAU1 channel.

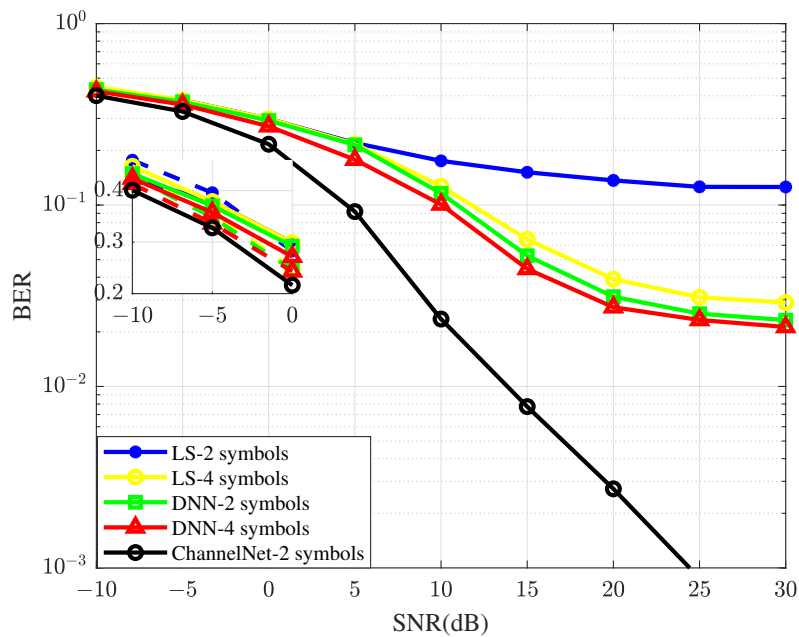


Figure 12. BER vs. SNR for NOF1 channel.

4.5. MSE and BER Performance for the Different Number of Pilot Symbols

Below, we analyzed the performance of MSE and BER under different pilot symbols. The above simulation used two symbols, occupying the 2nd and 14th OFDM symbols, respectively. The following situation combined the four pilot symbols, to occupy the 2nd, 6th, 10th, and 14th OFDM symbols, respectively. We first analyzed the case of MSE, as shown in Figures 13 and 14. It can be seen that because the input-channel pattern features corresponding to the added pilot symbols were richer, the performance of using four pilot symbols rather than two pilot symbols was significantly better in the high SNR. In the low-SNR case, the results of the LS-4 symbols were inferior to the LS-2-symbols case. The reason for this result is mainly attributable to the elements in the UWA-CSI matrix averaging in the order of 10^{-4} after the square, while the MSE in these cases was also of the same order. When the error is compared to itself, it is not surprising that anyone's MSE is better than the others, as their performance was disappointing.

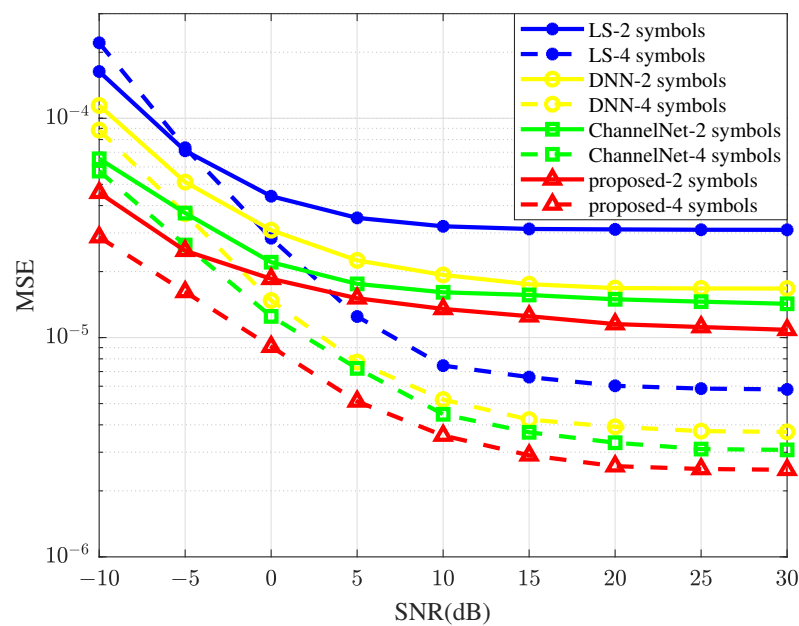


Figure 13. MSE vs. SNR for different number of pilot symbols in KAU1 channel.

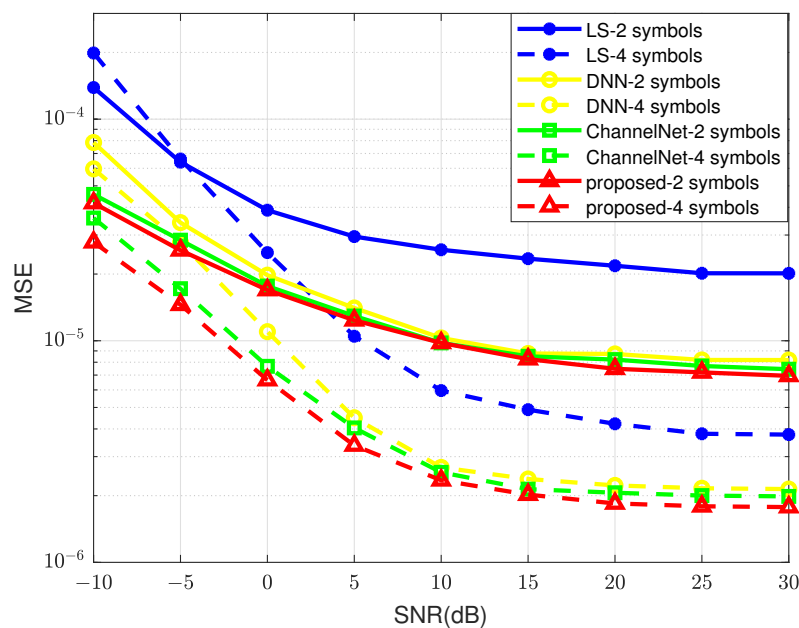


Figure 14. MSE vs. SNR for different number of pilot symbols in NOF1 channel.

The analysis of BER under different pilot symbols is shown in Figures 15 and 16. We can see that compared to the case of two symbols, four symbols can bring very excellent BER performance. It can also be seen from the figure that, whether in the KAU1 or NOF1 scenarios, our proposed algorithm can still maintain excellent performance, even with fewer pilot symbols. This result holds significant implications for the UWA-OFDM system within the repository. By conserving time-frequency resources for data transmission, we can notably enhance data rates.

We performed ablation experiments for the proposed CAD module, to demonstrate the effectiveness of the module. Our method took into account the uncertainty of the underwater environment for the channel estimation and did not directly estimate the channel information, but instead used the attention mechanism to denoise the channel. The experimental results are shown in Figure 17. We considered the effect of w/- or w/o CAD modules on the model under two different pilot signals. With the pilot signal of two and the low SNR = 5 dB, w/CAD improved the model performance by about 130% compared to w/o CAD. With the pilot signal of four and a high SNR = 25 dB, w/CAD was about 109% better than w/o CAD. This proves that CAD modules can further optimize channel information, and can selectively focus on or reinforce the information of the specific water-acoustic channel, especially in a low-SNR environment.

From the above simulation results, we know that the normalization amplitude of the channel estimated by our proposed algorithm was close to the actual channel. The performance of MSE and BER was better than the comparison algorithm, and it had better performance for underwater-acoustic channels of different complexity, especially for complex multi-path-structure scenarios in the performance comparison. For the small number of pilots, the proposed channel-estimation algorithm could also maintain better performance and help to enhance the data rate. However, because the samples of the experiment came from the enhanced samples generated by the CWGAN-CP model, which was close to the characteristics of the real samples, but had fine ambiguity compared to the real samples, the models trained in this paper may still have had a weak gap for the models trained with real data.

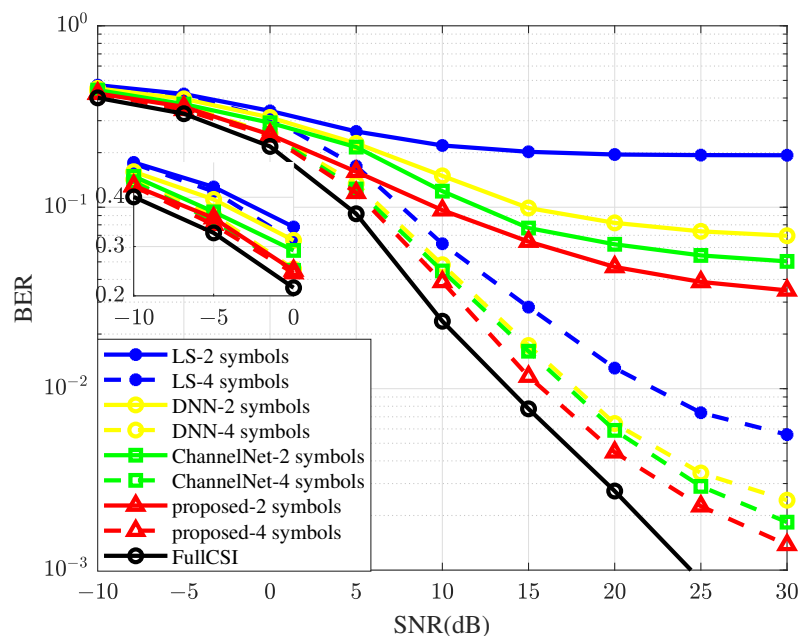


Figure 15. BER vs. SNR for different number of pilot symbols in KAU1 channel.

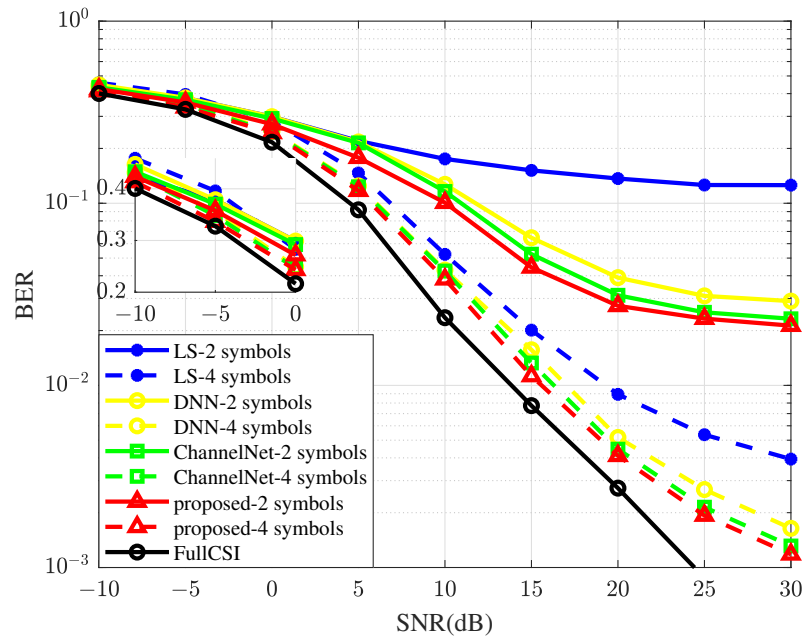


Figure 16. BER vs. SNR for different number of pilot symbols in NOF1 channel.

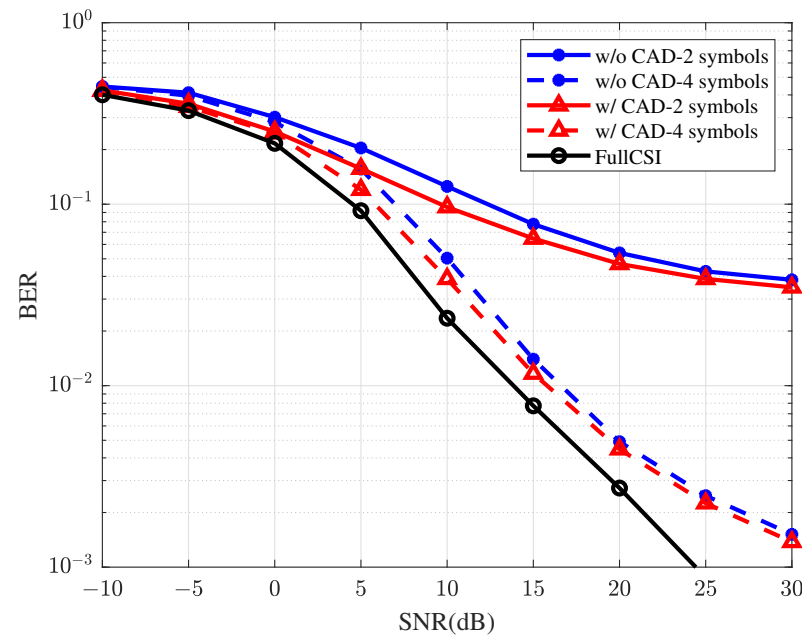


Figure 17. BER vs. SNR for different number of pilot symbols in CAD ablation experiments.

5. Conclusions

In this paper, we proposed a novel DL-based method for UWA-OFDM channel estimation. First, the CWGAN-GP model was introduced, to generate enhanced channel samples. Then, to reduce the influence of feature-information loss, we built a similar U-Net convolution-network structure to extract channel features in a skip-connected manner, then designed a channel-attention-denoising module, to further optimize the channel information reconstruction. Finally, the channel quality estimated by the proposed method was verified in KAU1 and NOF1 environments for different pilot symbols. The experimental results show that the channel response estimated by the proposed algorithm could fit the real channel samples well, and that the proposed algorithm was superior to the comparison algorithm in MSE and BER. For the low SNR, the proposed algorithm also achieved better performance with less pilot symbols, which could better improve the system-transmission performance. However, mathematical validation of the proposed

algorithm was difficult, due to the difficulty of deep-learning models. In the future, we will introduce more environmental data, such as number of users and number of channels, into the simulation validation.

Author Contributions: Conceptualization, J.G. and T.G.; methodology, J.G.; software, J.G.; validation, J.G. and M.L.; formal analysis, J.G. and H.L.; investigation, J.G. and M.L.; resources, T.G.; data curation, J.G.; writing—original draft preparation, J.G.; writing—review and editing, J.G. and T.W.; visualization, J.G.; supervision, T.G.; project administration, J.G. and T.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Wuzhou Science Plan Research Project, grant number 202202037, and by the Key Research Doctoral Fund Project of Wuzhou University, grant number 2022A004.

Data Availability Statement: Data is contained within the article.

Acknowledgments: The authors extend their heartfelt gratitude to the editors and reviewers for their insightful feedback and recommendations. The team is also deeply thankful for the invaluable contributions and advice provided by esteemed colleagues from the School of Computer, Electronics, and Information at Guangxi University and Wuzhou University. This work was supported by Guangxi Key Laboratory of Machine Vision and Intelligent Control, Wuzhou University, Wuzhou, Guangxi, China.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Pranitha, B.; Anjaneyulu, L. Review of research trends in underwater communications—A technical survey. In Proceedings of the 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 6–8 April 2016; pp. 1443–1447.
2. Singer, A.C.; Nelson, J.K.; Kozat, S.S. Signal processing for underwater acoustic communications. *IEEE Commun. Mag.* **2009**, *47*, 90–96. [[CrossRef](#)]
3. JIA Ning, H.J.C. An overview of underwater acoustic communications. *Physics* **2014**, *43*, 650–657.
4. Barazideh, R.; Niknam, S.; Natarajan, B.; V. Nikitin, A. Intermittently Nonlinear Impulsive Noise Mitigation and Doppler Shift Compensation in UWA-OFDM Systems. *IEEE Access* **2019**, *7*, 36590–36599. [[CrossRef](#)]
5. Li, B.; Zhou, S.; Stojanovic, M.; Freitag, L.; Willett, P. Multicarrier Communication Over Underwater Acoustic Channels with Nonuniform Doppler Shifts. *IEEE J. Ocean. Eng.* **2008**, *33*, 198–209.
6. Qiao, G.; Babar, Z.; Ma, L.; Liu, S.; Wu, J. MIMO-OFDM underwater acoustic communication systems—A review. *Phys. Commun.* **2017**, *23*, 56–64. [[CrossRef](#)]
7. Huang, J.; Wang, H.; He, C.; Zhang, Q.; Jing, L. Underwater acoustic communication and the general performance evaluation criteria. *Front. Inf. Technol. Electron. Eng.* **2018**, *19*, 951–971. [[CrossRef](#)]
8. Luo, H.; Wang, J.; Bu, F.; Ruby, R.; Wu, K.; Guo, Z. Recent Progress of Air/Water Cross-Boundary Communications for Underwater Sensor Networks: A Review. *IEEE Sensors J.* **2022**, *22*, 8360–8382. [[CrossRef](#)]
9. Tadayon, S.; Stojanovic, M. Iterative sparse channel estimation for acoustic OFDM systems. In Proceedings of the 2016 IEEE Third Underwater Communications and Networking Conference (UComms), Lercici, Italy, 30 August–1 September 2016; pp. 1–5.
10. Chen, Z.; He, Z.; Niu, K.; Rong, Y. Neural Network-based Symbol Detection in High-speed OFDM Underwater Acoustic Communication. In Proceedings of the 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP), Hangzhou, China, 18–20 October 2018; pp. 1–5.
11. Ozdemir, M.K.; Arslan, H. Channel estimation for wireless OFDM systems. *IEEE Commun. Surv. Tutor.* **2007**, *9*, 18–48. [[CrossRef](#)]
12. Altabbaa, M.T.; Ogrenici, A.S.; Panayirci, E.; Poor, H.V. Sparse Channel Estimation for Space-Time Block Coded OFDM-Based Underwater Acoustic Channels. In Proceedings of the 2018 IEEE Global Communications Conference (GLOBECOM), Abu Dhabi, United Arab Emirates, 9–13 December 2018; pp. 1–6.
13. Zhang, Y.; Li, J.; Zakharov, Y.; Li, X.; Li, J. Deep learning based underwater acoustic OFDM communications. *Appl. Acoust.* **2019**, *154*, 53–58. [[CrossRef](#)]
14. Zhang, J.; Cao, Y.; Han, G.; Fu, X. Deep neural network-based underwater OFDM receiver. *IET Commun.* **2019**, *13*, 1998–2002. [[CrossRef](#)]
15. Lee-Leon, A.; Yuen, C.; Herremans, D. Underwater Acoustic Communication Receiver Using Deep Belief Network. *IEEE Trans. Commun.* **2021**, *69*, 3698–3708. [[CrossRef](#)]
16. Jiang, R.; Wang, X.; Cao, S.; Zhao, J.; Li, X. Deep Neural Networks for Channel Estimation in Underwater Acoustic OFDM Systems. *IEEE Access* **2019**, *7*, 23579–23594. [[CrossRef](#)]

17. van Walree, P.A.; Jenserud, T.; Smedsrud, M. A Discrete-Time Channel Simulator Driven by Measured Scattering Functions. *IEEE J. Sel. Areas Commun.* **2008**, *26*, 1628–1637. [[CrossRef](#)]
18. Socheleau, F.X.; Laot, C.; Passerieux, J.M. Stochastic Replay of Non-WSSUS Underwater Acoustic Communication Channels Recorded at Sea. *IEEE Trans. Signal Process.* **2011**, *59*, 4838–4849. [[CrossRef](#)]
19. Huang, N.E.; Shen, Z.; Long, S.R.; Wu, M.C.; Shih, H.H.; Zheng, Q.; Yen, N.C.; Tung, C.C.; Liu, H.H. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A* **1998**, *454*, 903–998. [[CrossRef](#)]
20. Zhao, H.; Ji, F.; Li, Q.; Guan, Q.; Wang, S.; Wen, M. Federated Meta-Learning Enhanced Acoustic Radio Cooperative Framework for Ocean of Things. *IEEE J. Sel. Top. Signal Process.* **2022**, *16*, 474–486. [[CrossRef](#)]
21. Goodfellow, I.J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative Adversarial Nets. In Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, QC, Canada, 8–13 December 2014; Volume 2, pp. 2672–2680.
22. Creswell, A.; White, T.; Dumoulin, V.; Arulkumaran, K.; Sengupta, B.; Bharath, A.A. Generative Adversarial Networks: An Overview. *IEEE Signal Process. Mag.* **2018**, *35*, 53–65. [[CrossRef](#)]
23. Stojanovic, M.; Beaujean, P.P.J. Acoustic Communication. In *Springer Handbook of Ocean Engineering*; Springer: New York, NY, USA, 2016; pp. 359–386.
24. Adler, J.; Lunz, S. Banach wasserstein gan. *Adv. Neural Inf. Process. Syst.* **2018**, *31*. [[CrossRef](#)]
25. Gulrajani, I.; Ahmed, F.; Arjovsky, M.; Dumoulin, V.; Courville, A.C. Improved training of wasserstein gans. *Adv. Neural Inf. Process. Syst.* **2017**, *30*. [[CrossRef](#)]
26. Maas, A.L.; Hannun, A.Y.; Ng, A.Y. Rectifier nonlinearities improve neural network acoustic models. In Proceedings of the ICML 2013: 30th International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; Volume 30, p. 3.
27. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
28. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
29. Qarabaqi, P.; Stojanovic, M. Statistical Characterization and Computationally Efficient Modeling of a Class of Underwater Acoustic Communication Channels. *IEEE J. Ocean. Eng.* **2013**, *38*, 701–717. [[CrossRef](#)]
30. Soltani, M.; Pourahmadi, V.; Mirzaei, A.; Sheikhzadeh, H. Deep Learning-Based Channel Estimation. *IEEE Commun. Lett.* **2019**, *23*, 652–655. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.