Article

Deep Learning-Based Detection Algorithm for the Multi-User MIMO-NOMA System

Qixing Wang [1,2], Ting Zhou [3,4], Hanzhong Zhang [1], Honglin Hu [1], Edison Pignaton de Freitas [5] and Songlin Feng [1,*]

1   Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai 201210, China; wangqx@sari.ac.cn (Q.W.); zhanghz@sari.ac.cn (H.Z.); huhl@sari.ac.cn (H.H.)
2   School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 100049, China
3   School of Microelectronics, Shanghai University, Shanghai 200444, China; zhouting@shu.edu.cn
4   Shanghai Frontier Innovation Research Institute, Shanghai 201100, China
5   Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre 93950-000, Brazil; epfreitas@inf.ufrgs.br
*   Correspondence: fengsl@sari.ac.cn

Abstract: Recently, non-orthogonal multiple access (NOMA) has become prevalent in 5G communication. However, the traditional successive interference cancellation (SIC) receivers for NOMA still encounter challenges. The near-far effect between the users and the base stations (BS) results in a higher bit error rate (BER) for the SIC receiver. Additionally, the linear detection algorithm used in each SIC stage fails to eliminate the interference and is susceptible to error propagation. Consequently, designing a high-performance NOMA system receiver is a crucial challenge in NOMA research and particularly in signal detection. Focusing on the signal detection of the receiver in the NOMA system, the main work is as follows. (1) This thesis leverages the strengths of deep neural networks (DNNs) for nonlinear detection and incorporates the low computational complexity of the successive interference cancellation (SIC) structure. The proposed solution introduces a feedback deep neural network (FDNN) receiver to replace the SIC in signal detection. By employing a deep neural network for nonlinear detection at each stage, the receiver mitigates error propagation, lowers the BER in NOMA systems, and enhances resistance against inter-user interference (IUI). (2) We describe its algorithm flow and provide simulation results comparing FDNN and SIC receivers under MIMO-NOMA scenarios. The simulations clearly demonstrate that FDNN receivers outperform SIC receivers in terms of BER for MIMO-NOMA systems.

Keywords: NOMA; DNN receiver; multi-user system; MIMO

1. Introduction

The fifth generation (5G) of wireless communication technology has been widely applied to various types of services, including mobile Internet and Internet of Things (IoT) devices [1]. 5G technology has notably enhanced the transmission rates, minimized latency, and bolstered system capacity, enabling varied applications [2]. However, with the explosive growth in the number of mobile network users, the data traffic has increased 1000-fold by 2020 [3], and spectrum and time slot resources are limited [4]. The conventional orthogonal multiple access (OMA) systems face resource shortages with growing user access. To achieve the requirements for processing explosive data traffic, NOMA technology has been proposed. In a NOMA system, multiple users share the same frequency and time with distinct power allocations, thereby increasing user connections and system capacity. The earliest NOMA technology was proposed by NTT Corporation in Japan [5]. It can improve spectrum efficiency and increase system users' capacity. Recent studies have indicated that NOMA technology outperforms OMA in BER with the same bandwidth [6].

Another technology known as multiple-input multiple-output (MIMO) enables receivers with multiple antennas to communicate with multiple transmitters concurrently. In a massive MIMO scenario, terminals are equipped with numerous antennas, operating in a time division duplex, which improves throughput and radiated energy efficiency by concentrating energy into smaller spatial regions [7]. The emitted wavefronts add to the expected position and reduce their strength elsewhere. As a result, spatial multiplexing in massive MIMO significantly boosts the capacity of BS by several orders of magnitude [8].

In the MIMO-NOMA scenario, many scholars have conducted extensive research. In [9], a cooperative relay NOMA with a massive MIMO system based on the geometric-based stochastic channel model is proposed. In [10], hybrid precoding of millimeter wave large-scale MIMO based on NOMA is presented, which utilizes NOMA with beamforming to reduce the orthogonality error. Reference [11] incorporates code reuse in the proposed signal transmission scheme to access the BER of a MIMO-NOMA system. Reference [12] extends the millimeter wave mMIMO-NOMA algorithm to multiple user scenarios, and an SSOR-CRZF linear precoder is proposed. Reference [13] proposes a downlink NOMA algorithm for the 6G IoT scenario, ensuring users at different distances receive the same frequency signals from the base station. Reference [14] proposed an algorithm for distributing high-speed cached content using unmanned aerial vehicles (UAVs). Ground users can utilize these UAVs as airborne base stations, thereby minimizing the overall time latency of data transmission.

While NOMA technology is efficient in wireless communication, it faces certain drawbacks [15]. The computational complexity of NOMA system receivers poses high demands on receiver hardware and software performance. Additionally, the near-far effect between users and BS leads to higher BER in traditional receivers. Effectively designing a receiver system for NOMA has become a significant challenge. To address this, we integrate deep learning algorithms into the traditional receiver model. These algorithms leverage hidden layers in neural networks to train datasets, providing convergent solutions and proving efficiency in nonlinear detection scenarios. Much research has explored the combination of NOMA and deep learning algorithms. Reference [16] provides an overview of multiple NOMA systems supporting deep learning and proposes performance evaluation metrics. Reference [17] has explored user grouping for NOMA systems and applies the deep learning algorithm to power allocation. Reference [18] utilizes deep neural networks to enhance performance in power optimization for a downlink NOMA system, leveraging the long short-term memory (LSTM) network to predict channel coefficients. Reference [19] proposes a DNN-based algorithm that combines reconfigurable intelligent surfaces (RISs) to enhance predictability and security performance. Reference [20] applies the DNN network to power allocation algorithms, which can improve the performance of the downlink visible light communication NOMA system in factory automation scenarios. Reference [21] proposed an algorithm applying simultaneously transmitting and reflecting RIS (STAR-RISs) in MIMO-NOMA scenarios. The algorithm jointly optimizes the reflection, transmission phase shifts, and precoding matrix of the STAR-RIS based on each user's Quality of Service (QoS). Additionally, a convolutional neural network (CNN) model is utilized to predict the downlink spectral efficiency by analyzing user location information and channel gains.

One research direction in this field involves applying deep learning methods to the receivers or detectors for MIMO-NOMA systems. Reference [22] introduces a machine-learning-based MIMO signal detector algorithm that can learn joint detection without requiring channel state information (CSI). Reference [23] proposed a DNN-based joint detector for the orthogonal frequency-division multiplexing (OFDM)-based NOMA scheme, performing channel estimation, equalization, and demodulation simultaneously. Reference [24] applies a CNN network to the scene of a single BS and multi-user downlink NOMA receivers. The convolutional layer enhances feature extraction, reducing errors introduced during the SIC detection process. Reference [25] suggests an LSTM deep learning network employing nonlinear mapping algorithms to assess data detection capacity in NOMA systems, ensuring high accuracy with low complexity. Reference [26] uses DNN

to construct the precoder and SIC decoders. This approach enables the precise decoding of received signals at users through the SIC decoding technique after precoding at the transmitter. Reference [27] applies two deep learning models to the MIMO-NOMA visible light communication (MIMO-NOMA-VLC) scenario. Simulation results demonstrate that the models can achieve low BER in both perfect and imperfect SIC scenarios. Reference [28] utilizes a deep learning model with a discrete wavelet transform (DWT)-based OFDM approach. By analyzing CSI and channel noise information, this algorithm can reduce inter-channel interference.

In this article, a DNN-based SIC receiver algorithm for the MIMO-NOMA scene is proposed. We apply the DNN method to traditional SIC receivers, which can reduce the errors caused by the near-far effect and enhance nonlinear detection capability. We discuss the selection of the DNN network structure and the appropriate number of hidden layers. Simulation results demonstrate the impact of the signal-to-noise ratio (SNR) on BER. Performance comparisons with traditional SIC receivers affirm the effectiveness of the DNN-based receiver algorithm.

In the rest of the paper, Section 2 introduces the system model. Section 3 shows the proposed method. Section 4 presents the extensive experiments. Finally, Section 5 will summarize the conclusions.

2. System Model

2.1. NOMA System

NOMA facilitates the simultaneous transmission of symbols by multiple users on the same frequency. Power domain and code domain multiplexing are two distinct types of NOMA. Focusing on power domain multiplexing, it enables the user's signal to be superimposed and encoded in the power domain, allowing a resource block to serve multiple users concurrently. Considering an uplink NOMA scheme with N users, the channel gain for the i-th user is h_i. Assuming the channel gain from user 1 to user N is arranged in descending order, which is $h_1 > h_2 > \cdots > h_n$, the power allocation p_i should be $P_1 > P_2 > \cdots > P_n$. It is because the difference in user symbol gain should be more pronounced to make the station distinguish signals accurately as it receives the user's superimposed signal. Consequently, users with stronger channel gains are allocated higher power. The superimposed signal y at the BS can be written as [29]

$$y = \sum_{i=1}^{N} \sqrt{P_i} h_i s_i + w_0 \,.$$ (1)

In the SIC receiver, signal detection occurs based on the user's symbol gain, progressing from larger gains to smaller gains. When detecting users with strong h_i, the receiver treats the signals of users with weak h_i as interference signals.

In the uplink NOMA scenario, the SINR of the i-th user can be calculated by the superimposed signals from the base station receiver, which is written as [30]

$$\text{SINR}_i = \frac{P_i |h_i|^2}{\sum_{k=i+1}^{N} P_k |h_k|^2 + N_0} \,.$$ (2)

Assuming the bandwidth of the transmission time-frequency block is B, the throughput R_i of the i-th user in the uplink NOMA system can be written as

$$R_i = B \log \left(1 + \frac{P_i |h_i|^2}{\sum_{k=i+1}^{N} P_k |h_k|^2 + N_0} \right) \,.$$ (3)

Assuming the transmission power of two users is fixed at 1 W, the bandwidth is 1 Hz, the channel noise at the base station end is 10 mW, the power distribution factor of user 1 and user 2 is α and $1 - \alpha$. Due to the higher h_i of user 1 than user 2, the base station receiver

first detects the symbols of user 1 and then detects the symbols of user 2. At this point, the throughput of users 1 and 2 in the uplink NOMA scenario can be represented as [6]

$$R_1^{NOMA} = \log\left(1 + \frac{\alpha P|h_1|^2}{(1-\alpha)P|h_2|^2 + N_0}\right). \tag{4}$$

$$R_2^{NOMA} = \log\left(1 + \frac{(1-\alpha)P|h_2|^2}{N_2}\right). \tag{5}$$

2.2. SIC Receiver

The fundamental concept of the SIC receiver is to detect corresponding symbols in descending order based on signal amplitude. Therefore, the SIC receiver initially sorts the amplitude of each user's signal during the detection process. In the uplink NOMA scenario, the BS detects symbols for each user in descending order, considering the product of the channel gain and power for user i. Typically, each stage of the SIC receiver employs a linear detection algorithm such as zero-forcing (ZF) [31] detection. Assuming the received signal is Y, the symbol gain matrix is H, the actual symbol sent by the sending end is X, and the estimated symbol at the receiving end is X. The objective function of the ZF detection algorithm is $\min L(x) = |Y - H\hat{X}|_F^2$, in which

$$\begin{aligned} L(x) &= |Y - H\hat{X}|_F^2 = (Y - H\hat{X})^H(Y - H\hat{X}) \\ &= Y^H Y - Y^H H\hat{X} - \hat{X}^H H^H Y + \hat{X}^H H^H H\hat{X}. \end{aligned} \tag{6}$$

The derivative form of the objective function can be denoted as

$$\frac{d(L(X))}{d\hat{X}} = -2H^H Y + 2H^H H\hat{X} = 0. \tag{7}$$

Then, we can establish the linear relationship between the received symbols and the symbol estimation at the receiving end. Specifically, the weight matrix of the ZF detection algorithm can be performed as

$$\hat{X} = (H^H H)^{-1} H^H Y = W_{ZF} Y. \tag{8}$$

Assuming the original superimposed signal at the receiving terminal is Y, the symbol gain for the i-th user is H_i, and the estimated symbol value for the i-th user is $\hat{s}_i$. In the n-th level detection, the superimposed signal can be denoted as

$$Y^{(n)} = Y - \sum_{i=1}^{n-1} H_i \hat{s}_i, 1 \le n \le N. \tag{9}$$

Through N-level symbol detection and interference cancellation, the SIC receiver can obtain symbol estimates for all N users.

The SIC receiver boasts low computational complexity. However, relying on a linear detection method, it may encounter error propagation issues in case of an error. In the uplink NOMA system, the constrained transmission power of terminals makes it challenging to achieve significant power differences through user matching and power control strategies, leading to interference between users. Additionally, the near-far effect between users and the BS in the uplink scenario exacerbates the situation. The signals of these users are superimposed on the same resource block, potentially resulting in irregular superimposed constellations. Consequently, due to the linear detection algorithm used by the SIC receiver at each detection level, errors may arise in detecting signals within irregular constellation diagrams. As a result, the BER can be significantly higher than that of the downlink scenario. To address these shortcomings, we propose the application of a DNN on the SIC receiver.

2.3. DNN Network

The deep learning method has found widespread applications in various fields such as natural language processing, image recognition, and the Internet of Things. Figure 1 illustrates the basic structure of a DNN network, which can be described as three parts. The input layers receive the input data matrix from different users and transmit it to the hidden layers. With interconnected neuron nodes in the hidden layers, the DNN network enhances accuracy in data identification. As the data enter the hidden layer, the function expression can be denoted as

$$z_j = \sum_{i=1}^{n} w_{ij} x_i + b_j \,, \tag{10}$$

in which x_i denotes the i-th node of the input layer, and b_i denotes the threshold value of the node in the hidden layer. Between two different layers, w_{ij} represents the weight vector. Since the neural network introduces the nonlinear activation function, the DNN structure can be used for nonlinear mapping. In this paper, we choose the Relu function as the activation function. Then, the output of the j-th node can be denoted as

$$O_j = f(z_j) \,. \tag{11}$$

By introducing the Relu function in each hidden layer, the final output can be expressed as

$$\hat{y} = f(a_j) = f(\sum_{j=1}^{m} w_{ji} O_j + b_i) \,, \tag{12}$$

where a_j denotes the j-th node in the hidden layer which is directly connected with the output layer.
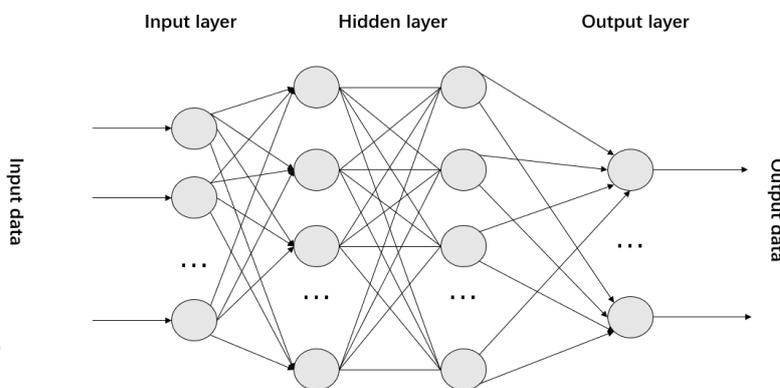

Figure 1. Structure of the DNN model.

By introducing a DNN algorithm, its nonlinear activation function and hidden layer structure can enable the receivers to have the ability of nonlinear detection, enhancing detection accuracy and reducing errors. Details about the selection of the activation function, the number of neuron nodes, and the hidden layers in the DNN network model used in this paper are provided in Sections 3.2.2 and 4.2.

3. Proposed Method

This section introduces a feedback deep neural networks-based algorithm for MIMO-NOMA detection. Initially, we establish the fundamental structure of the MIMO-NOMA system and subsequently present the FDNN algorithm designed to address signal detection challenges.

3.1. MIMO-NOMA System

Consider an uplink MIMO-NOMA system. In this system, a BS equipped with N antennas is employed. To mitigate near-field effects, users are required to maintain a certain distance from the BS. The BS receives signals simultaneously from K users, each transmitting with a single antenna, which can be written as

$$\tilde{y} = \tilde{H}\tilde{x} + \tilde{n}, \tag{13}$$

where $\tilde{x}$ represents the user's modulation symbols vector. $\tilde{H}$ represents the channel matrix during transmission, and $\tilde{n}$ represents the AWGN noise. We transform (13) into a real domain as

$$y = Hx + n, \tag{14}$$

where $H \triangleq \begin{bmatrix} \mathcal{R}(\tilde{H}) & -\mathcal{S}(\tilde{H}) \\ \mathcal{S}(\tilde{H}) & \mathcal{R}(\tilde{H}) \end{bmatrix}$, $y \triangleq \begin{bmatrix} \mathcal{R}(\tilde{y}) \\ \mathcal{S}(\tilde{y}) \end{bmatrix}$, $x \triangleq \begin{bmatrix} \mathcal{R}(\tilde{x}) \\ \mathcal{S}(\tilde{x}) \end{bmatrix}$, $n \triangleq \begin{bmatrix} \mathcal{R}(\tilde{n}) \\ \mathcal{S}(\tilde{n}) \end{bmatrix}$. The real parts are written as $\mathcal{R}(\cdot)$ and the imaginary parts are written as $\mathcal{S}(\cdot)$. We have $H \in \mathbb{R}^{2N \times 2K}$, $y \in \mathbb{R}^{2N}$, $n \in \mathbb{R}^{2N}$ and $x \in \mathbb{R}^{2K}$. In this system, K users share the total power P, which can be divided into $p_1, p_2, \cdots, p_k$. The power allocation coefficient for user *k* is denoted as $\alpha_k = \frac{p_k}{P}$, where $\sum_{k=1}^{K} \alpha_k = 1$. By monitoring the CSI during transmission, users can be arranged in ascending order based on the magnitude of the channel gain h_k. Users with higher channel gain are assigned greater α_k, enhancing signal clarity during reception. However, α_k cannot increase without limitations. Allocating excessive power to one user may distort the information of other users during transmission, leading to an overall increase in errors. Balancing efficiency and equity considerations, the power allocation factor should fluctuate within a certain interval. Detailed simulation results are presented in Section 4.4.

To enhance transmission rates, reduce IUI, and lower BER, numerous researchers have explored user-matching strategies. In the downlink NOMA scenario, the BS possesses higher transmit power and greater computational capabilities. Utilizing user-matching strategies in the downlink can create larger power differences among NOMA signals in the same resource block, effectively reducing IUI. However, in the uplink NOMA system, terminal transmit power is constrained, making it challenging to achieve significant power differences through user-matching strategies.

3.2. FDNN Model

Detecting symbols sent by users in a communication system essentially involves a judgment process akin to a multi-categorization problem. So, we employ one-hot coding to represent different classifications and leverage feedback deep neural networks for non-linear classification. Using one-hot coding helps reduce computational complexity when calculating the cross-entropy of the vectors, which facilitates the training and testing of the network. The FDNN receiver follows a detection sequence from strong to weak user signals. After each detection level, the reconstructed signal is fed back to the input terminal and subtracted from the superimposed signal. Subsequently, the detection process for the next user ensues. The FDNN receiver model is depicted in Figure 2.

The received signal y and the channel parameter vector Θ serve as the inputs of the FDNN network. They undergo multiple hidden layers, each comprising a fully connected layer followed by a ReLu layer. In the last hidden layer, the input from the preceding layer undergoes a fully connected layer and a softmax layer. The output then passes through a reconstruction layer, and this reconstructed signal is subtracted from y. The rest of it will become the new received signal for the subsequent user's FDNN receiver.
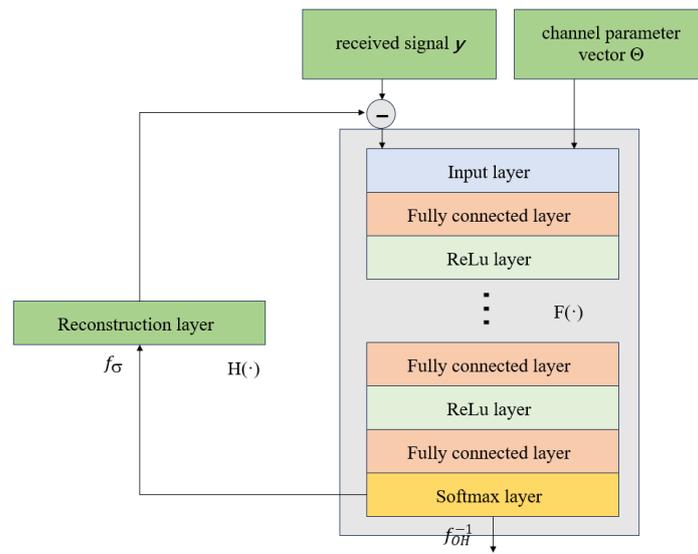
Figure 2. FDNN receiver model.

3.2.1. Data Preprocessing

The initial step involves updating the superimposed signals and channel parameters. The superimposed signal is updated before the detection process for each user. When detecting the i-th user, the superimposed signal y_i is computed as the sum of the original signal minus the reconstructed signal from the previous user, which is written as

$$y_i = y - \sum_{k=1}^{i} \mathrm{H}(\hat{x}_k),\tag{15}$$

y represents the original signal that the BS receives. $\mathrm{H}(\cdot)$ represents the reconstruction function, which is responsible for recovering the user's signal based on encoding estimation. With y_i, we can extract the in-phase component y_i^I and orthogonal component y_i^Q as the input of the network. The channel parameter vector can be expressed as

$$\boldsymbol{\Theta}_i = [g_i, \varphi_i, g_{i+1}, \varphi_{i+1}, \cdots, g_N, \varphi_N].\tag{16}$$

g_i is the symbol gain and $g_i = \sqrt{P_i}|h_i|$. φ_i is the symbol phase of user i at the receiving end and $\varphi_i = \arctan(\frac{Im(h_i)}{Re(h_i)})$. The channel parameter vector is updated before each user's detection by subtracting the channel parameter of the previous user. Therefore, the input vector of the FDNN network is $I_i = [y_i^I, y_i^Q; \boldsymbol{\Theta}_i]$. At the output terminal, we map the modulation symbol s_i of user i to the one-hot code x_i, which is written as

$$x_i = f_{OH}(s_i), s_i \in S_i.\tag{17}$$

The one-hot code makes it easy to calculate the cross-entropy of vectors.

3.2.2. Detection Module

The detection module encompasses a fully connected deep neural network structure featuring nonlinear layers. This neural network consists of K fully connected layers with the initial $K - 1$ fully connected layers directly succeeded by an activation layer. The ReLu function is selected as the activation function, which is written as

$$f_{ReLu}(x) = \max(0, x).\tag{18}$$

Additionally, we employ the Softmax classification function at the end of the FDNN network, where the output represents the probability of the estimated symbol. The Softmax function can be represented as

$$f_{Softmax}(x_n) = \frac{e^{x_n}}{\sum_m e^{x_m}} .$$ (19)

The total detection module can be expressed as

$$\begin{aligned} \hat{x}_i &= \mathrm{F}(I_i) \\ &= f_{Softmax}(W_K f_{ReLu}(W_{K-1} \cdots W_1 f_{ReLu}(I_i + b_1) \cdots + b_{K-1}) + b_K), \end{aligned}$$ (20)

in which W_K and b_K represent the weight matrix and bias vector of the k-th fully connected layer. $\hat{x}_i$ signifies the estimated value of the one-hot code output by the detection module. This estimated value can be transformed into the estimated modulation symbols $\hat{s}_i$ by

$$\hat{s}_i = f_{OH}^{-1}(\hat{x}_i), \hat{s}_i \in S_i .$$ (21)

3.2.3. Feedback Module

The feedback module is responsible for completing the reconstruction of the detection signal. The decision function f_σ binarizes the output vector into a unit vector. By multiplying the binary unit vector with the reconstruction gain matrix, the reconstruction signal can be obtained. The reconstruction module can be expressed as

$$\mathrm{H}(\hat{x}_i) = f_\sigma(\hat{x}_i) V_i + c_i ,$$ (22)

in which V_i and c_i represent the reconstruction gain matrix and reconstruction bias vector. Since the channel parameter information of each user is known, V_i and c_i can be constructed based on the channel parameter information and updated only based on the detected users. Typically, modulation symbols are unbiased, so the $c_i = 0$. V_i can be denoted as

$$V_i = [\sqrt{P_i} h_i s_i^{(1)}, \sqrt{P_i} h_i s_i^{(2)}, \cdots, \sqrt{P_i} h_i s_i^{(M)}]^T ,$$ (23)

in which P_i and h_i are the power and the channel parameter for user i. $s_i^{(1)}, s_i^{(2)}, \cdots, s_i^{(M)}$ denoted the M modulation symbols of user i.

3.2.4. Model Training

The FDNN receiver transforms symbol detection into a classification problem. We chose Kullback–Leibler (KL) divergence to measure the distribution differences of vectors. The mathematical expression for solving the target can be expressed as

$$\min_{\mathrm{F}(\cdot), \mathrm{F}(\cdot)} D_{KL}(x_i || \hat{x}_i) = \sum_{k=1}^{M} x_i^{(k)} \log \frac{x_i^{(k)}}{\hat{x}_i^{(k)}} ,$$ (24)

in which x_i and $\hat{x}_i$ represent the label values and estimated values of one-hot code for user i; $x_i^{(k)}$ represents the value of the k-th position of the one-hot coding vector of user i. M is the number of modulation symbols for user i. $\mathrm{F}(\cdot)$ and $\mathrm{H}(\cdot)$ are the function expressions of the detection module and the reconstruction module.

Next, we need to determine the loss function $L(\hat{x}, x)$. Considering that the training samples are known, the entropy of the training sample label $H(x_i)$ is a constant. Since we have $D(x_i || \hat{x}_i) = H(x_i, \hat{x}_i) - H(x_i)$, solving for the minimum of (24) is equivalent to solving for the minimum of the cross-entropy $H(x_i, \hat{x}_i)$. As the receiver aims to minimize the detection error for each user, the loss function is the sum of cross-entropy losses for each user's detection. The loss function can be written as

$$L(\hat{\boldsymbol{x}}, \boldsymbol{x}) = -\sum_{i=1}^{N} \sum_{k=1}^{M} x_i^{(k)} \log \hat{x}_i^{(k)} + \lambda \sum_k w_k^2. \tag{25}$$

The neural network employs stochastic gradient descent for training, representing the parameter update method for the detection module as

$$w_{mn}^k(t+1) = w_{mn}^k(t) - \alpha \nabla w_{mn}^k, \tag{26}$$

$$b_n^k(t+1) = b_n^k(t) - \alpha \nabla b_n^k, \tag{27}$$

in which $w_{mn}^k(t)$ represents the values of the weight matrix $\boldsymbol{W}_k$ in the m-th row and n-th column of the k-th layer during the t-th iterations. $b_n^k(t)$ represents the values of the bias vector $\boldsymbol{b}_k$ in the n-th column of the k-th layer during the t-th iteration. $\alpha \nabla w_{mn}^k$ and $\alpha \nabla b_n^k$ are the gradient values and α is the learning rate of the neural network.

4. Simulation Results

In this section, we assess the influence of neural network hidden layer structure and training batch size on performance. Subsequently, we present BER performance under varying Δ_{SNR}dB between users. Finally, a performance comparison is made between FDNN receivers and SIC receivers, and the training parameters are detailed in Table 1.

Table 1. Training setting in numerical tests.

Parameters	Values
N	2
K	2
Modulation Mode	QPSK
Hidden Layers	2
Trainnig Sample Size	200,000
Regularization Parameter λ	0.01
Learning Rate α	0.02
Batch Size	100
Epoch	100

Consider a two-user uplink NOMA system using QPSK modulation. Due to varying distances between users and the BS, a near-far effect induces channel gain and phase differences. A learning rate α of 0.02 and regularization parameter λ of 0.01 were chosen for optimal performance after preliminary testing. The hidden layers, trainng sample size, and batch size will be discussed in the subsections below.

4.1. Data Set

We utilize MATLAB R2021b to produce communication signal data for both training and testing sets. Assuming a NOMA-MIMO system with two users and one BS equipped with two antennas, the signal data spans from SNR = 0 dB to SNR = 30 dB in 1 dB increments. With a block number of 2000 and a batch size of 100, the overall training sample size amounts to 200,000.

4.2. Hidden Layers

The hidden layer structure is depicted in Figure 2. Three types of hidden layers were compared. Type 1 features a three-layer structure with an input layer, two hidden layers, and an output layer, each with 16 and 32 nodes. Types 2 and 3 are five-layer structures with input and output layers, along with four hidden layers. The nodes in type 2 are (16, 32, 48, 32), while in type 3, they are (32, 128, 64, 20). Δ_{SNR} represents the difference in SNR between two user signals at the receiving end. When $\Delta_{SNR} = 5$ dB or $\Delta_{SNR} = 7$ dB, Figures 3 and 4 show the BER performance of them.
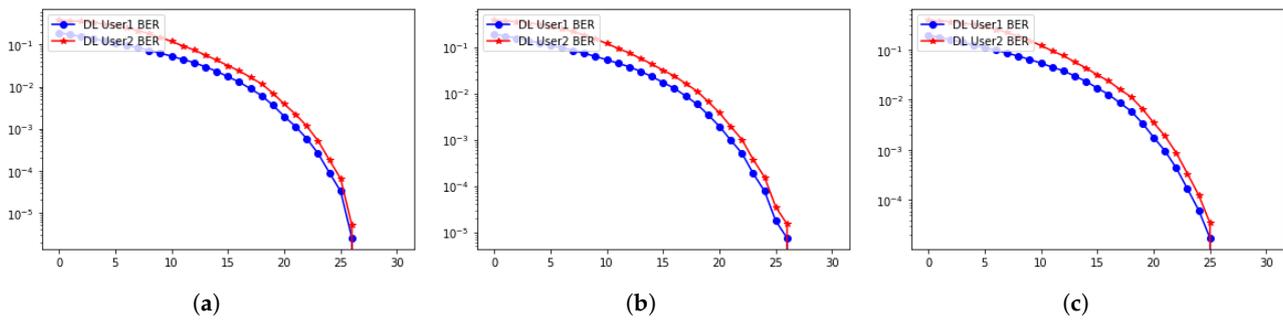
Figure 3. $\Delta_{SNR} = 5$ dB for three types: (**a**) type 1; (**b**) type 2; (**c**) type 3.
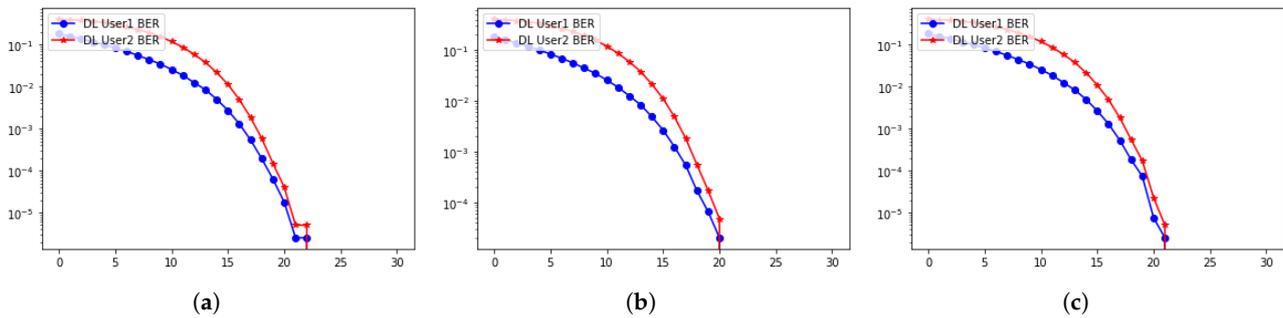

Figure 4. $\Delta_{SNR} = 7$ dB for three types: (**a**) type 1; (**b**) type 2; (**c**) type 3.

From Figures 3 and 4, networks with five hidden layers perform better than networks with three hidden layers. When the network structures are identical, those with more neuron nodes exhibit better performance. Nevertheless, the performance difference between three-layer and five-layer network structures is not significant. Considering computational complexity and energy consumption, we opt for the three-layer structure.

4.3. Batch Size

This article also simulates and analyzes the training speed of the FDNN model. We set the batch size as 10, 100, and 1000 symbols. Figure 5 shows the training loss curves under different batch sizes.
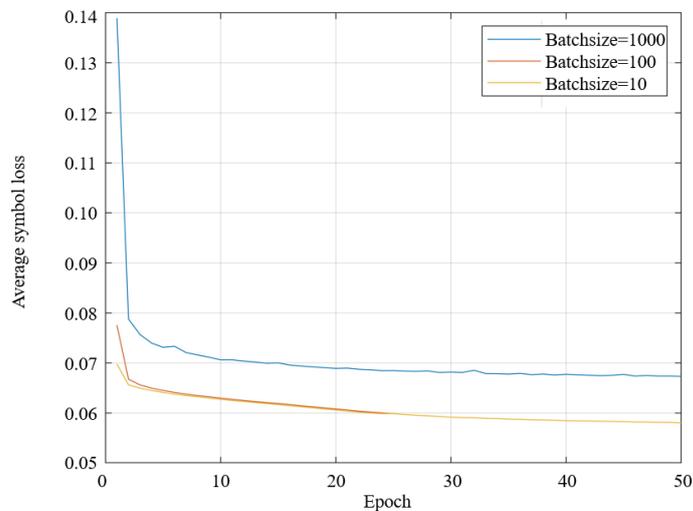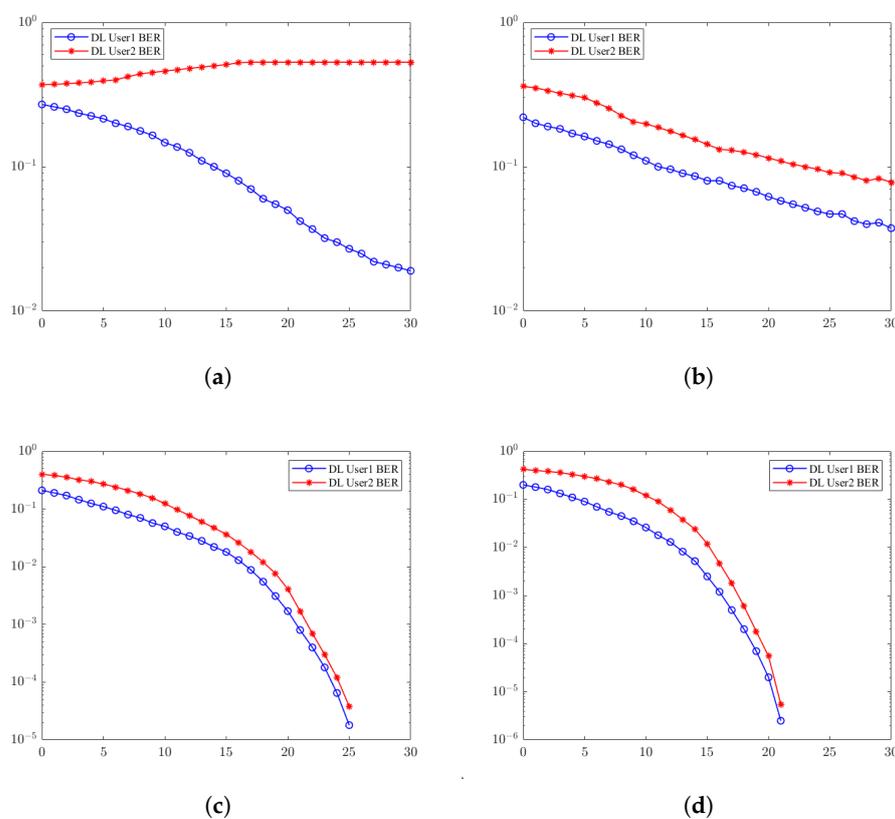

Figure 5. Performance of FDNN under different batch sizes.

With increasing training epochs, the training loss consistently decreases and converges to a small value. The convergence loss for batch sizes of 10 and 100 samples is notably lower than that for a batch size of 1000 samples. In earlier training rounds, the loss for a batch size of 10 samples decreased more rapidly than that for 1000 samples and slightly faster than 100 samples. However, considering computational complexity and the final convergence loss, a batch size of 100 samples is deemed more suitable.

Moreover, the training error rapidly decreases to a low value within the initial 10 rounds when employing a batch size of 100 samples, suggesting that the FDNN model can achieve quick convergence with lower computational demands. This characteristic is advantageous for practical applications where low latency is essential.

4.4. Performance of FDNN Receivers

Figure 6 shows the performance of FDNN receivers under different Δ_{SNR}.

Figure 6. Performance of FDNN receivers under different Δ_{SNR}. (**a**) $\Delta_{SNR} = 1$ dB; (**b**) $\Delta_{SNR} = 3$ dB; (**c**) $\Delta_{SNR} = 5$ dB; (**d**) $\Delta_{SNR} = 7$ dB.

When $\Delta_{SNR} = 1$ dB, user 2 struggles to reduce BER due to the small Δ_{SNR}, making it challenging for receivers to separate signals. So, the BER curve of user 2 is close to a parallel horizontal line without a downward trend. As the Δ_{SNR} increases, BER descent accelerates, indicating improved signal distinguishability. When $\Delta_{SNR} = 3$ dB, BER curves show a linear decline from 10^{-1} to 10^{-2}. Further increases in $\Delta_{SNR} = 5$ dB and $\Delta_{SNR} = 7$ dB lead to even faster BER reduction, reaching 10^{-4} for $\Delta_{SNR} = 5$ dB and 10^{-6} for $\Delta_{SNR} = 7$ dB at SNR $= 22$ dB. This demonstrates that a higher Δ_{SNR} correlates with a lower BER.

Figures 7 and 8 show the performance of FDNN and SIC receivers for user 1 and user 2.
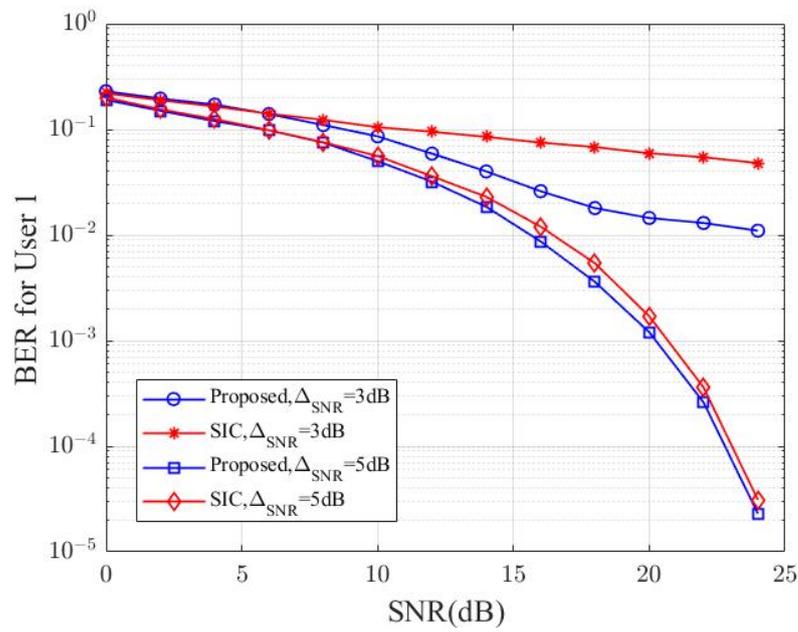
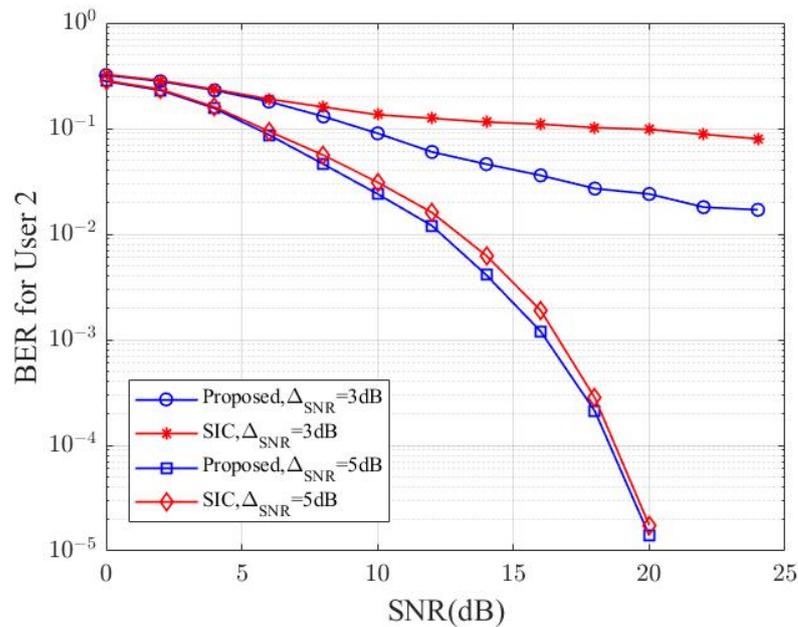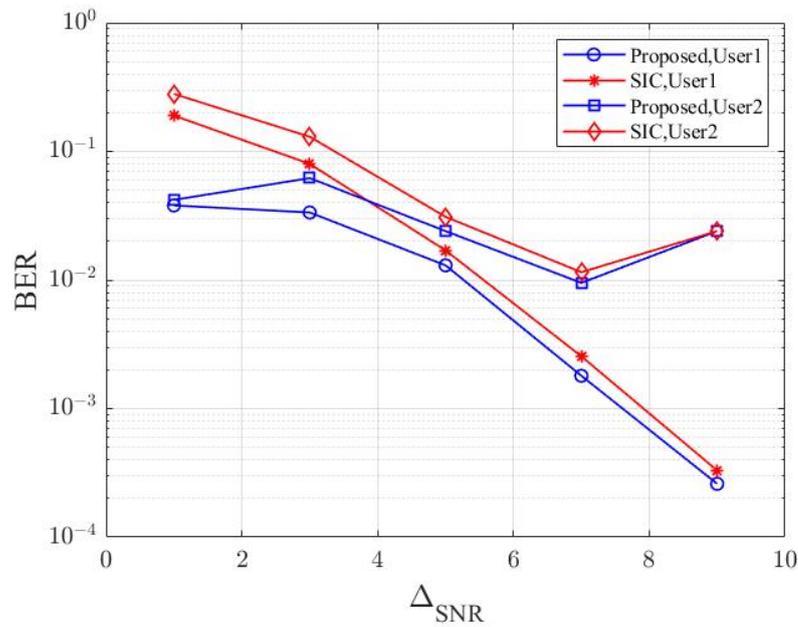**Figure 7.** Performance of FDNN and SIC receivers for user 1.

Figure 8. Performance of FDNN and SIC receivers for user 2.

Table 2 presents the BER for two users at SNR $=$ 20 dB. Both user 1 and user 2 exhibit lower BERs with FDNN receivers compared to SIC receivers. As user SNR increases, the BER advantage of FDNN receivers over SIC receivers becomes more pronounced. Additionally, the comparison between $\Delta_{SNR} = 3$ dB and $\Delta_{SNR} = 5$ dB reveals that FDNN receivers perform better than SIC receivers at $\Delta_{SNR} = 3$ dB, emphasizing the significant impact of SNR difference on NOMA user detection performance. FDNN receivers excel in scenarios with lower user differences.

Table 2. BER at SNR = 20 dB.

User	$\Delta_{SNR} = 3$ dB	$\Delta_{SNR} = 5$ dB
User 1 for SIC	5.95×10^{-2}	1.7×10^{-3}
User 1 for FDNN	1.45×10^{-2}	1.2×10^{-3}
User 2 for SIC	9.8×10^{-2}	1.75×10^{-5}
User 2 for FDNN	2.4×10^{-2}	1.4×10^{-5}

Figure 9 examines the influence of the SNR difference Δ_{SNR} between users on the BER of NOMA systems. The greater the SNR of a channel, the more power it is allocated. We use this to simulate the effect of power allocation on system performance.

Figure 9. Performance of FDNN and SIC receivers under different Δ_{SNR}.

Table 3 presents the BER for two users at different Δ_{SNR}. The *x*-axis represents Δ_{SNR} between two users, and the y-axis represents BER. User 1 maintains a fixed SNR of 15 dB, while the SNR of user 2 is $(15 - \Delta_{SNR})$dB. We can see that when $\Delta_{SNR} = 3 \sim 7$ dB, as Δ_{SNR} increases, the BER for both two users decreases due to reduced interference. However, when $\Delta_{SNR} = 9$ dB, the BER of user 1 continues to decline, while the BER for user 2 increases, as the SNR of user 2 becomes too low. The comparison between the blue and red lines in Figure 9 indicates that FDNN receivers consistently outperform SIC receivers, reinforcing the superiority of FDNN receivers in MIMO-NOMA scenarios.

Table 3. BER at different Δ_{SNR}.

User	$\Delta_{SNR} = 3$ dB	$\Delta_{SNR} = 7$ dB	$\Delta_{SNR} = 9$ dB
User 1 for SIC	8.03×10^{-2}	2.55×10^{-3}	3.37×10^{-4}
User 1 for FDNN	3.35×10^{-2}	1.8×10^{-3}	2.6×10^{-4}
User 2 for SIC	1.3×10^{-1}	1.15×10^{-2}	2.45×10^{-2}
User 2 for FDNN	6.22×10^{-2}	9.52×10^{-3}	2.46×10^{-2}

5. Conclusions

This paper proposed an FDNN-based approach for signal detection in MIMO-NOMA scenarios. This algorithm addresses the high BER issue in SIC receivers caused by the

near-far effect and error propagation. By leveraging the advantages of DNN in nonlinear detection and the low computational complexity of SIC, this paper proposes an FDNN model to replace the traditional SIC receiver, providing its algorithm flow. Simulation results demonstrate the algorithm's computational efficiency and effectiveness.

Based on simulation results, this paper chooses 100 as the training batch size, and the neural network has a three-layer structure consisting of an input layer, two hidden layers with node counts of 16 and 32, and an output layer. This configuration enables the FDNN model to achieve better performance. This paper assesses the FDNN algorithm's effectiveness and examines the influence of the power allocation scheme on system performance. Simulation outcomes consistently demonstrate that FDNN receivers exhibit lower BER compared to SIC receivers, affirming the performance advantages of FDNN in MIMO-NOMA scenarios. In future research, the performance gap of FDNN under different conditions and scenarios will be further investigated, aiming to enhance the algorithm's stability in engineering applications.

Author Contributions: Conceptualization, Q.W. and T.Z.; methodology, Q.W.; software, H.Z.; validation, H.H.; writing—original draft preparation, Q.W. and E.P.d.F.; writing—review and editing, T.Z.; supervision, S.F. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: Derived data supporting the findings of this study are available from the corresponding author on request.

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

References

1. Saad, W.; Bennis, M.; Chen, M. A Vision of 6G Wireless Systems: Applications, Trends, Technologies, and Open Research Problems. *IEEE Netw.* **2020**, *34*, 134–142. [CrossRef]
2. Zhang, H.; Zhou, T.; Xu, T.; Hu, H. Remote Interference Discrimination Testbed Employing AI Ensemble Algorithms for 6G TDD Networks. *Sensors* **2023**, *23*, 2264. [CrossRef] [PubMed]
3. Akyildiz, I.F.; Kak, A.; Nie, S.T. 6G and Beyond: The Future of Wireless Communications Systems. *IEEE Access* **2020**, *8*, 133995–134030. [CrossRef]
4. Wu, J.; Xu, T.; Zhou, T.; Chen, X.; Zhang, N.; Hu, H. Feature-Based Spectrum Sensing of NOMA System for Cognitive IoT Networks. *IEEE Internet Things J.* **2023**, *10*, 801–814. [CrossRef]
5. Saito, Y.; Kishiyama, Y.; Benjebbour, A.; Nakamura, T.; Li, A.; Higuchi, K. Non-Orthogonal Multiple Access (NOMA) for Cellular Future Radio Access. In Proceedings of the IEEE 77th Vehicular Technology Conference (VTC Spring), Dresden, Germany, 2–5 June 2013.
6. Ding, Z.; Lei, X.; Karagiannidis, G.K.; Schober, R.; Yuan, J.; Bhargava, V.K. A Survey on Non-Orthogonal Multiple Access for 5G Networks: Research Challenges and Future Trends. *IEEE J. Sel. Areas Commun.* **2017**, *35*, 2181–2195. [CrossRef]
7. Larsson, E.G.; Edfors, O.; Tufvesson, F.; Marzetta, T.L. Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **2014**, *52*, 186–195. [CrossRef]
8. Agiwal, M.; Roy, A.; Saxena, N. Next Generation 5G Wireless Networks: A Comprehensive Survey. *IEEE Commun. Surv. Tutor.* **2016**, *18*, 1617–1655. [CrossRef]
9. Tweneboah-Koduah, S.; Affum, E.A.; Agyekum, K.A.P.; Ajagbe, S.A.; Adigun, M.O. Performance of Cooperative Relay NOMA with Large Antenna Transmitters. *Electronics* **2022**, *11*, 3482. [CrossRef]
10. Alraddady, F.; Ahmed, I.; Habtemicail, F. Robust Hybrid Beam-Forming for Non-Orthogonal Multiple Access in Massive MIMO Downlink. *Electronics* **2022**, *11*, 75. [CrossRef]
11. Gkonis, P.K.; Trakadas, P.T.; Sarakis, L.E. Non-Orthogonal Multiple Access in Multiuser MIMO Configurations via Code Reuse and Principal Component Analysis. *Electronics* **2020**, *9*, 1330. [CrossRef]
12. Sur, S.N.; Kandar, D.; Silva, A.; Nguyen, N.D.; Nandi, S.; Do, D.T. Hybrid Precoding Algorithm for Millimeter-Wave Massive MIMO-NOMA Systems. *Electronics* **2022**, *11*, 2198. [CrossRef]
13. Xie, W.L.; Ding, X.; Cai, B.W.; Li, X.; Wei, M.S. Downlink MIMO-NOMA System for 6G Internet of Things. *Electronics* **2022**, *11*, 3233. [CrossRef]
14. El-Gayar, M.M.; Ajour, M.N. Resource Allocation in UAV-Enabled NOMA Networks for Enhanced Six-G Communications Systems. *Electronics* **2023**, *12*, 5033. [CrossRef]

15. Dai, L.L.; Wang, B.C.; Yuan, Y.F.; Han, S.F.; Chih-Lin, I.; Wang, Z.C. Non-Orthogonal Multiple Access for 5G: Solutions, Challenges, Opportunities, and Future Research Trends. *IEEE Commun. Mag.* **2015**, *53*, 74–81. [CrossRef]

16. Mohsan, S.A.H.; Li, Y.; Shvetsov, A.V.; Varela-Aldás, J.; Mostafa, S.M.; Elfikky, A. A Survey of Deep Learning Based NOMA: State of the Art, Key Aspects, Open Challenges and Future Trends. *Sensors* **2023**, *23*, 2946. [CrossRef]

17. Li, J.; Gao, T.; He, B.; Zheng, W.; Lin, F. Power Allocation and User Grouping for NOMA Downlink Systems. *Appl. Sci.* **2023**, *13*, 2452. [CrossRef]

18. Gaballa, M.; Abbod, M.; Aldallal, A. Investigating the Combination of Deep Learning for Channel Estimation and Power Optimization in a Non-Orthogonal Multiple Access System. *Sensors* **2022**, *22*, 3666. [CrossRef]

19. Dang, H.P.; Nguyen, M.S.V.; Do, D.T.; Nguyen, M.H.; Pham, M.T.; Kim, A.T. Secure Performance Analysis of Aerial RIS-NOMA-Aided Systems: Deep Neural Network Approach. *Electronics* **2022**, *11*, 2588. [CrossRef]

20. Ryu, W.J.; Kim, J.W.; Kim, D.S. Resource Allocation in Downlink VLC-NOMA Systems for Factory Automation Scenario. *Sensors* **2022**, *22*, 9407. [CrossRef]

21. Perdana, R.H.Y.; Nguyen, T.-V.; An, B. A Deep Learning-Based Spectral Efficiency Maximization in Multiple Users Multiple STAR-RISs Massive MIMO-NOMA Networks. In Proceedings of the 2023 Fourteenth International Conference on Ubiquitous and Future Networks (ICUFN), Paris, France, 4–7 July 2023.

22. Shlezinger, N.; Fu, R.; Eldar, Y.C. DeepSIC: Deep Soft Interference Cancellation for Multiuser MIMO Detection. *IEEE Trans. Wirel. Commun.* **2021**, *20*, 1349–1362. [CrossRef]

23. Chen, W.; Tang, Z. Research on improved receiver of NOMA-OFDM signal based on deep learning. In Proceedings of the 2021 International Conference on Communications, Information System and Computer Engineering (CISCE), Beijing, China, 14–16 May 2021.

24. Sim, I.; Sun, Y.G.; Lee, D.; Kim, S.H.; Lee, J.; Kim, J.H.; Shin, Y.; Kim, J.Y. Deep Learning Based Successive Interference Cancellation Scheme in Nonorthogonal Multiple Access Downlink Network. *Energies* **2020**, *13*, 6237. [CrossRef]

25. Gui, G.; Huang, H.J.; Song, Y.W.; Sari, H. Deep Learning for an Effective Nonorthogonal Multiple Access Scheme. *IEEE Trans. Veh. Technol.* **2018**, *67*, 8440–8450. [CrossRef]

26. Kang, J.M.; Kim, I.; Chun, C.J. Deep Learning-Based MIMO-NOMA with Imperfect SIC Decoding. *IEEE Syst. J.* **2020**, *14*, 3414–3417. [CrossRef]

27. Salama, W.M.; Aly, M.H.; Amer, E.S. Deep learning based BER improvement for NOMA-VLC systems with perfect and imperfect successive interference cancellation. *Opt. Quantum Electron.* **2023**, *55*, 692. [CrossRef]

28. Ahmad, M.; Shin, S.Y. Wavelet-based massive MIMO-NOMA with advanced channel estimation and detection powered by deep learning. *Phys. Commun.* **2023**, *61*, 102189. [CrossRef]

29. Islam, S.M.R.; Avazov, N.; Dobre, O.A.; Kwak, K.S. Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges. *IEEE Commun. Surv. Tutorials* **2017**, *19*, 721–742. [CrossRef]

30. Dai, L.L.; Wang, B.C.; Ding, Z.G.; Wang, Z.C.; Chen, S.; Hanzo, L. A Survey of Non-Orthogonal Multiple Access for 5G. *IEEE Commun. Surv. Tutor.* **2018**, *20*, 2294–2323. [CrossRef]

31. Ding, Z.G.; Adachi, F.; Poor, H.V. The Application of MIMO to Non-Orthogonal Multiple Access. *IEEE Trans. Wirel. Commun.* **2016**, *15*, 537–552. [CrossRef]