

Article

Deep Convolutional and Recurrent Neural-Network-Based Optimal Decoding for RIS-Assisted MIMO Communication

Md Habibur Rahman ^{1,2}, Mohammad Abrar Shakil Sejan ^{1,2}, Md Abdul Aziz ^{1,2}, Dong-Sun Kim ³, Young-Hwan You ^{2,4} and Hyoung-Kyu Song ^{1,2,*}

¹ Department of Information and Communication Engineering, Sejong University, Seoul 05006, Republic of Korea; habibur@sju.ac.kr (M.H.R.); sejan@sejong.ac.kr (M.A.S.S.); aziz@sju.ac.kr (M.A.A.)

² Department of Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, Republic of Korea; yhyou@sejong.ac.kr

³ Department of Semiconductor Systems Engineering, Sejong University, Seoul 05006, Republic of Korea; dongsun.kim07@sejong.ac.kr

⁴ Department of Computer Engineering, Sejong University, Seoul 05006, Republic of Korea

* Correspondence: songhk@sejong.ac.kr

Abstract: The reconfigurable intelligent surface (RIS) is one of the most innovative and revolutionary technologies for increasing the effectiveness of wireless systems. Deep learning (DL) is a promising method that can enhance system efficacy using powerful tools in RIS-based environments. However, the lack of extensive training of the DL model results in the reduced prediction of feature information and performance failure. Hence, to address the issues, in this paper, a combined DL-based optimal decoding model is proposed to improve the transmission error rate and enhance the overall efficiency of the RIS-assisted multiple-input multiple-output communication system. The proposed DL model is comprised of a 1-dimensional convolutional neural network (1-D CNN) and a gated recurrent unit (GRU) module where the 1-D CNN model is employed for the extraction of features from the received signal with further process over the configuration of different layers. Thereafter, the processed data are used by the GRU module for successively retrieving the transmission signal with a minimal error rate and accelerating the convergence rate. It is initially trained offline using created OFDM data sets, after which it is used online to track the channel and extract the transmitted data. The simulation results show that the proposed network performs better than the other technique that was previously used in terms of bit error rate and symbol error rate. The outcomes of the model demonstrate the suitability of the proposed model for use with the next-generation wireless communication system.

Keywords: reconfigurable intelligent surface (RIS); 1-dimensional convolutional neural network (1-D CNN); gated recurrent unit (GRU); bit-error rate (BER); symbol-error rate (SER)

MSC: 94A14



Citation: Rahman, M.H.; Sejan, M.A.S.; Aziz, M.A.; Kim, D.-S.; You, Y.-H.; Song, H.-K. Deep Convolutional and Recurrent Neural-Network-Based Optimal Decoding for RIS-Assisted MIMO Communication. *Mathematics* **2023**, *11*, 3397. <https://doi.org/10.3390/math11153397>

Academic Editors: Farag M. Sallabi, Mohammad Hayajneh and Daniel-Ioan Curiac

Received: 12 June 2023

Revised: 19 July 2023

Accepted: 2 August 2023

Published: 3 August 2023



Copyright: © 2023 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

The growth of the fifth-generation (5G) mobile communication has caused a significant bandwidth demand. In wireless networks, greater frequency bands are allotted and used to enhance higher data rates and system capacity. The 5G New Radio (5G NR) is given to mmWave frequencies between 24 GHz and 50 GHz. This technical development is expected to lead to the employment of the sub-THz bands of 114–300 GHz in wireless communication systems beyond the (B5G) or 6G systems [1]. For the next B5G or 6G wireless networks, the reconfigurable intelligent surface (RIS) is being studied as a viable new technology [2]. The real-time programmable controller in RIS is controlled by a number of passive devices that may independently alter the phase of electromagnetic wave propagation [3,4]. The RIS can be implemented in the walls of the inside and outside environment of the buildings and

windows glasses [5–7]. The usage of RIS might be used to enhance wireless communications by removing channel blockages, boosting non-line-of-sight connections, expanding the coverage, and lowering inter-user interference (IUI). As each of its components has the capacity to alter the phase and amplitude of the incoming signal, RIS may offer passive beam-forming (BF) [8] by modifying the phase shift and amplitude reflection coefficients of its constituent parts. In contrast to the conventional multiple-input multiple-output (MIMO) systems, which place a greater emphasis on the base station (BS) and user BF, the RIS-assisted communication system requires the joint design of active BF for BS and user, as well as passive BF for RIS, in order to achieve the passive BF gain. Additionally, RIS-assisted systems enhance spectral efficiency (SE) and energy efficiency (EE) while using less money and little power. This is implemented by employing large-scale passive reflecting components [9–11]. The literature has recently been devoted to several research activities covering various aspects, such as channel modeling [12], channel (CE) estimation [13], modulation and encoding [14], SE analysis [6,15], outage probability [16], symbol error probability (SER) [17], energy efficiency [18], weighted sum rate [19,20], and the evaluation of performance in RIS-supported wireless networks [21]. This is because RIS-enabled settings provide significant potential advantages.

Deep learning (DL) approaches have the potential to significantly improve wireless communication performance [22–26]. Many works were performed in RIS-based wireless systems to handle problems related to channel state information (CSI) [27], CE [28], and performance maximization [29,30] using the DL approaches. To avoid performance loss using a DL model, in [29], the authors proposed an end-to-end training-based RIS-assisted MIMO system. The proposed system optimized the signal processing operations at the access point, RIS, and user at the same time, with active BF for the access point and user and passive BF for RIS. For estimating and identifying symbols in signals sent by RIS, the authors of [30] designed a DL method that employs fully connected layers for prediction when estimating channels and phase angles from a reflected signal received through a RIS. In [31], the authors proposed an effective RIS-assisted MIMO system based on a DNN multi-stage training strategy model for perfect CSI acquisition. The authors in [32] proposed a deep denoising neural-network-assisted compressive sensing broadband for the mmWave RIS system. An indoor RIS-based communication environment was considered in [33], where a trained DNN was used to determine the optimal phase shift by using the target user position. A deep reinforcement learning framework was proposed in [34] for predicting the RIS reflection matrix with minimal beam training overhead.

For natural language processing and machine vision, the convolutional neural network (CNN) model is basically used. It can extract the high level of features from the input and also can be used for the time-series data processing [35]. The CNN model is composed of an input layer, convolutional layer, pooling layer, fully connected layer, and output layer. The structure of the 1-D CNN is illustrated in Figure 1a.

In [36], the authors proposed a CNN-based RIS-enhanced multiple-input single-output system for the goal of sum-rate maximization. To archive the accurate CSI and improve performance, an ordinary differential equation-based CNN structured in the RIS-orthogonal frequency division multiplexing (OFDM) single antenna receiver was proposed in [37]. In [38], a CNN-based RIS-assisted OFDM with a single-input single-output system was proposed for achieving the maximization of the achievable rate. In [39], the authors presented a long short-term memory (LSTM)-based energy-efficient RIS wireless system where they evaluated the system EE and achievable rate. The study in [40] presented a CNN-based demodulation method for a multi-user RIS wireless system where, for channel modeling, an OFDM system based on MIMO was taken into consideration. In [40], the authors compared the achievement of the proposed demodulation (Demod)-CNN with the conventional method in terms of bit-error rate (BER) and SER. In addition, the temporal neural network (TCN)-based RIS-assisted MIMO wireless communication was presented in [41], where the proposed TCN analyzed the achievable rate in terms of different modulation schemes.

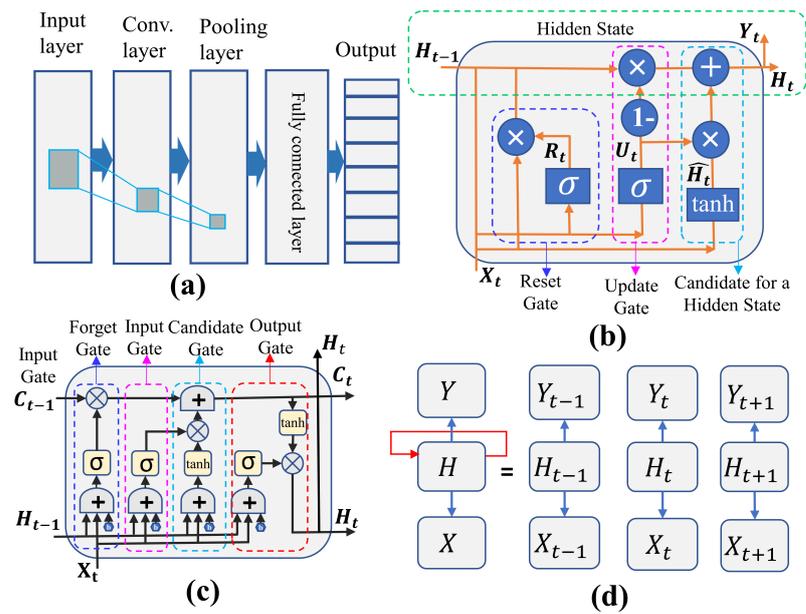


Figure 1. (a) The architecture of the 1-D CNN model; (b) the architecture of the GRU model; (c) the architecture of the LSTM model; (d) the architecture of the RNN model.

LSTM and gated recurrent unit (GRU) are examples of recurrent neural networks (RNNs) that are frequently employed to address sequence difficulties. In comparison to traditional RNN, GRU is designed to store past state information more effectively. The traditional RNN training process’s gradient vanishing and gradient explosion problems can be addressed [42]. The hidden unit in the typical RNN structure, which may selectively recall relevant information and discard unnecessary information, is replaced by a gate structure in the GRU and LSTM. The update gate and the reset gate are used in the GRU instead of the input gate, forget gate, and output gate of the LSTM [43]. The training setting can be decreased to produce a quicker convergence speed, provided that the prediction accuracy of the GRU is not less than that of the LSTM. Figure 1b shows the internal cell structure of the GRU network. Figure 1c,d show the structure of LSTM and typical RNN network, respectively. For a system with a one-bit ADC and massive MIMO, in [44], the authors proposed the LSTM-GRU to evaluate the channel matrix and showed better performance. For the purpose of CE of data subcarriers, in [45], the authors proposed a DL-GRU neural network-based OFDM system in which the computing demands may be greatly decreased and achieved a more satisfactory gain than other methods.

The performance in the aforementioned studies has a major influence on the SER and BER by using CNN- and RNN-based models. Thus, the combination of 1-D CNN with GRU can be a promising method to achieve high accuracy performance in RIS-assisted communication systems. Motivated by the previous work and enhanced efficacy, herein, a combined DL model, comprising a 1-D CNN and GRU network, is proposed. In terms of SER and BER with various signal-to-noise (SNR) levels, the suggested model outperforms the [40].

The key contributions of the proposed system can be summarized as follows:

- DL-cascaded model-based effective decoding is proposed for the RIS-assisted MIMO system. The proposed DL model is structured with a 1-D CNN model, which extracts the features in the form of complex numbers of the received channel matrix, and a GRU network, which uses extracted data for obtaining the optimal prediction.
- Inter-carrier interference and inter-symbol interference (ISI) are taken into consideration while segmenting the data of the received signals using a 1-D CNN-based feature extractor. For handling the considerable ISI and accelerating the training speed of the model, the GRU layers are utilized. Thus, the combined model enhances demodulation accuracy.

- The different layers of the proposed model are configured for improved accuracy and effectiveness and are examined by an Adam optimizer.
- The efficiency of the proposed model is compared with a conventional system, the LSTM model implemented in our experiment, TCN [41], and (Demod)-CNN-based RIS model [40], with respect to the different SNR values where the proposed model outperforms other methods at different SNRs with the measurement error of BER and SER.

The remaining parts of the paper are arranged as follows. The system model for RIS-based communication, which includes the channel model and signal transmission, is described in Section 2. The training data preparation, training, and inference of the proposed deep learning model are covered in Section 2.4. The simulation outcomes and complexity analysis are reported in Section 3, and Section 4 presents the conclusion.

Notations: A vector and matrix are denoted by the lowercase and uppercase bold letters \mathbf{h} and \mathbf{H} , respectively. \mathbf{H}^H stands for the conjugate transpose matrix of \mathbf{H} , and $\text{diag}(\mathbf{x})$ stands for the diagonal matrix with vector \mathbf{x} on its diagonal.

2. RIS-Assisted System and Channel Model

In this section, the RIS-assisted system, channel modeling, and selected precoder structure are described.

2.1. System Model

Figure 2 depicts a RIS communication system where I user equipment (UE) contains a single antenna, and the base station (BS) contains M uniform linear array (ULA) antennas. The total number of RIS is taken to be uniform planner array (UPA) N elements. For the i -th UE ($i = 1, 2, \dots, I$), the received signal through RIS can be stated as follows [6]:

$$\mathbf{y}_i = \mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br} \mathbf{x} + \mathbf{n}_i, \tag{1}$$

where $\mathbf{x} \in \mathbb{C}^{M \times 1}$ represents the precoded transmit signal with Q subcarriers in each OFDM symbol. Inverse discrete Fourier transform (IDFT) is used to process the OFDM data symbols after they are transformed to the time domain. Then, the CP data are inserted as a guard interval. Discrete Fourier transform (DFT) is carried out at the UE after the CP data are removed. Thereafter, \mathbf{y}_i is the signal that is received at the i -th UE. $\mathbf{H}_{br} \in \mathbb{C}^{N \times M}$ is the channel matrix from BS to RIS, the channel matrix between RIS and the i -th UE is $\mathbf{H}_{2,i} \in \mathbb{C}^{N \times 1}$, and at the i -th user, the additive white Gaussian noise is represented $n_i \sim \mathcal{CN}(0, \sigma^2)$. The phase shift values of RIS elements are represented by $\mathbf{\Omega} = \text{diag}(\mathbf{\Psi}) \in \mathbb{C}^{N \times N}$, which is the diagonal matrix of $\mathbf{\Omega}$. Each element can be characterized as $\mathbf{\Psi} = [\gamma_1 e^{j\zeta_1}, \gamma_2 e^{j\zeta_2}, \dots, \gamma_N e^{j\zeta_N}]^T \in \mathbb{C}^{N \times 1}$, where the phase shift coefficient and the amplitude for the n -th reflective elements are $\zeta_n \in [0, 2\pi]$ and $\gamma_n \in [0, 1]$. Consequently, the symbol $\mathbf{\Omega}$ can be written as follows:

$$\mathbf{\Omega} = \begin{bmatrix} \gamma_1 e^{j\zeta_1} & \dots & \dots & \dots \\ \dots & \gamma_2 e^{j\zeta_2} & \dots & \dots \\ \vdots & \vdots & \dots & \vdots \\ \dots & \dots & \dots & \gamma_N e^{j\zeta_N} \end{bmatrix}. \tag{2}$$

The constant amplitude coefficient is assumed to be $\zeta_n = 1$ for convenience in calculations [9]. In this study, the direct communication channel $\mathbf{H}_d \in \mathbb{C}^{M \times 1}$ between BS to UEs is also considered. The BS to the i -th UEs channel $\mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br} \in \mathbb{C}^{M \times 1}$ via RIS is assumed. In particular, it is important to remember that the matrix $\mathbf{\Omega} = \text{diag}(\mathbf{\Psi})$ is a diagonal matrix. Thereafter, the previous whole channel matrix $\mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br}$ can be reformed as follows:

$$\mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br} = \mathbf{H}_{2,i}^H \text{diag}(\mathbf{\Psi}) \mathbf{H}_{br} = \mathbf{\Psi}^T \text{diag}(\mathbf{H}_{2,i}^H) \mathbf{H}_{br}, \tag{3}$$

$$\mathbf{H}_{cas} = \mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br}, \tag{4}$$

where the cascaded channel for the i -th UEs is $\mathbf{H}_{cas} \in \mathbb{C}^{N \times M}$ and only relies on the downlink CSI [46]. Therefore, the total received signal \mathbf{y}_{rit} at the i -th UEs via the cascaded channel and the direct communication link can be written as follows [47]:

$$\mathbf{y}_{rit} = \mathbf{H}_{2,i}^H \mathbf{\Omega} \mathbf{H}_{br} \mathbf{x} + \mathbf{H}_d \mathbf{x} + \mathbf{n}_i, \tag{5}$$

where \mathbf{H}_d is the direct communication link from BS to the i -th UEs.

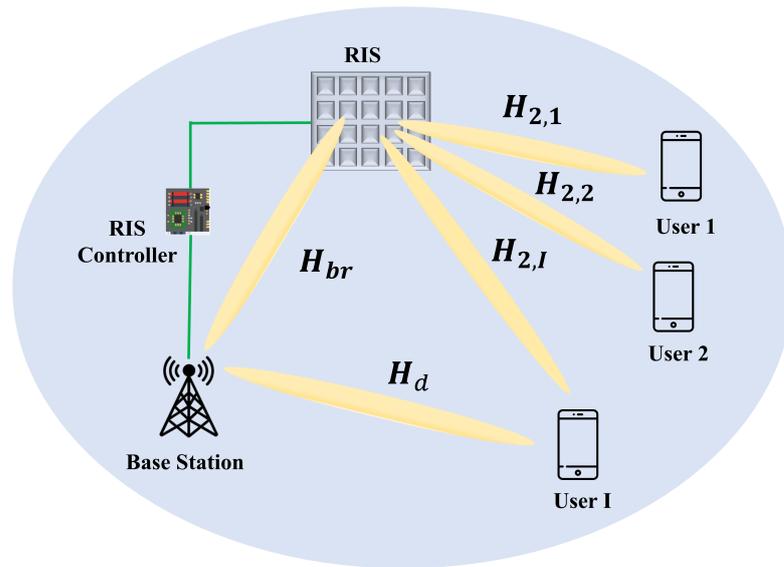


Figure 2. The architecture of RIS-assisted downlink MIMO communication.

2.2. Channel Model

The three-dimensional (3D) Saleh-Valenzuela channel model is used in this research for the mmWave propagation [48], which is a statistical channel model in a multipath propagation scenario. The prevalent theory of the channel model for mmWave communication can be represented as follows [49]:

$$\mathbf{H} = \sqrt{\frac{N}{L}} \sum_{l=1}^L \eta_l \mathbf{a}(\zeta_l^H, \phi_l^H), \tag{6}$$

where the channel vector is represented as \mathbf{H} , the complex gain of the l -th path is η_l , and the number of total paths is L . The azimuth angle of departure and the elevation angle of departure are ζ_l^H and ϕ_l^H , respectively. The array response vector is defined as $\mathbf{a}(\zeta_l^H, \phi_l^H)$. The array response vector for a standard $N_1 \times N_2$ is expressed as follows [49]:

$$\mathbf{a}(\zeta, \phi) = \frac{1}{\sqrt{N}} [e^{-j2\pi d \sin(\zeta) \cos(\phi) \mathbf{n}_1 / \lambda} \otimes e^{-j2\pi d \sin(\phi) \mathbf{n}_2 / \lambda}], \tag{7}$$

where $\mathbf{n}_1 = [0, 1, \dots, N_1 - 1]$ and $\mathbf{n}_2 = [0, 1, \dots, N_2 - 1]$, the carrier wavelength is λ , and the antenna spacing is d , the antenna inter-space satisfies the condition $d = \lambda/2$. Accordingly, the BS to RIS channel \mathbf{H}_{br} can be presented as follows:

$$\mathbf{H}_{br} = \sqrt{\frac{MN}{L_1}} \sum_{l_1=1}^{L_1} \eta_{l_1} \mathbf{b}(\zeta_{l_1}^{H_r}, \phi_{l_1}^{H_r}) \mathbf{a}^H(\zeta_{l_1}^{H_t}, \phi_{l_1}^{H_t})^T, \tag{8}$$

where L_1 stands for the pathway numbers among the BS and RIS, η_l stands for the complex gain of those paths, $\mathbf{b}(\zeta_l^{H_r}, \phi_l^{H_r})$ is the steering vector related to the RIS, and $\mathbf{a}(\zeta_l^{H_t}, \phi_l^{H_t})$ is the steering vector connected to the BS for the l -th path.

The following definitions can be used to describe the channel $\mathbf{H}_{2,i}^H$ between the RIS and the UE, as follows:

$$\mathbf{H}_{2,i}^H = \sqrt{\frac{N}{L_2}} \sum_{l_2=1}^{L_2} \eta_{l_2} \mathbf{a}^H(\zeta_{l_2}^{H_t}, \phi_{l_2}^{H_t}), \tag{9}$$

where the complex gain of paths is η_{l_2} , and L_2 is the number of pathways among the RIS and the i -th UE. $\zeta_{l_2}^{H_t}$ and $\phi_{l_2}^{H_t}$ represent the departure angles in azimuth and elevation, respectively. The array response vector is defined as $\mathbf{a}(\zeta_{l_2}^{H_t}, \phi_{l_2}^{H_t})$. In conclusion, according to (4), (8), and (9), the cascaded channel matrix \mathbf{H}_{cas} from BS to the i -th UE can be expressed as follows [6]:

$$\mathbf{H}_{cas} = \sqrt{\frac{MN}{L_1 L_2}} \sum_{l_1=1}^{L_1} \sum_{l_2=2}^{L_2} \eta_{l_1} \eta_{l_2} \text{diag}(\mathbf{a}^H(\zeta_{l_2}^{H_t}, \phi_{l_2}^{H_t})) \mathbf{b}(\zeta_{l_1}^{H_r}, \phi_{l_1}^{H_r}) \mathbf{a}^H(\zeta_{l_1}^{H_t}, \phi_{l_1}^{H_t})^T. \tag{10}$$

2.3. Design of Precoder

Dirty paper coding (DPC), successive interference cancellation (SIC), maximum ratio transmission (MRT), zero-forcing (ZF), and minimal mean squared error (MMSE) are examples of linear and non-linear precoders that can be taken into consideration in a MIMO system. Although non-linear precoders exhibit the best performance, implementation is impractical because of the enormous computational complexity. Linear precoders can be implemented in the MIMO system due to the low complexity with sub-optimal performance. In multi-user systems, the linear precoder of ZF performs IUI removal more efficiently than the linear precoder of MRT. Additionally, the ZF precoder performs similarly to the MMSE precoder when the noise value is taken into account [50]. Hence, ZF precoder $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_I] \in \mathbb{C}^{M \times I}$ is implemented in this study. For total matrix channel $\mathbf{H}_{all} = [\mathbf{H}_1, \dots, \mathbf{H}_I]^T \in \mathbb{C}^{I \times M}$, the precoder \mathbf{Z} can be written as follows:

$$\mathbf{Z} = \bar{\mathbf{Z}} \sqrt{\mathbf{P}}, \tag{11}$$

where the ZF precoder matrix stands for $\bar{\mathbf{Z}}$, and the power allocation matrix is represented by $\mathbf{P} = \text{diag}(\delta_1, \dots, \delta_I)$. The i -th UE's transmit power at the BS is represented by the δ_i symbol.

2.4. Proposed Deep Learning Model

The dataset generation, structure, and operation of the model are described, and the training and testing procedures are described at the end of this section.

2.4.1. Dataset Preparation

In this paper, 128 OFDM subcarriers are taken into consideration for the transmission of data, and quadrature phase shift-keying (QPSK) modulation is used to map the bits belonging to symbols. The data are transmitted over a RIS-assisted cascaded and direct channel. The BS sends the whole OFDM packet to the UE along with the additional AWGN noise. The generated feature vector \mathbf{V}_{ce} and matching label are used to store the received OFDM packet as training data samples. The number of labels is created by the combination of data from BS antennas where 4 unique symbols can be generated by each antenna; in this way, $4^2 = 16$ labels are generated where 2 is the BS antennas number. The feature vector \mathbf{V}_{ce} holds the complex signal of each symbol in the OFDM packet. To train the suggested model, the complex signal is first divided into real \mathfrak{R}_c and imaginary \mathfrak{I}_m components. The generated transmission and reception data are divided into training ξ and validation τ

data, respectively. A total of 256,000 data samples are generated where 80% are used for training ξ and 20% for validation τ . Finally, the dataset and label samples are stored for the training and validation process. The example datasets of the real and imaginary parts of symbols with corresponding labels are shown in Figure 3.

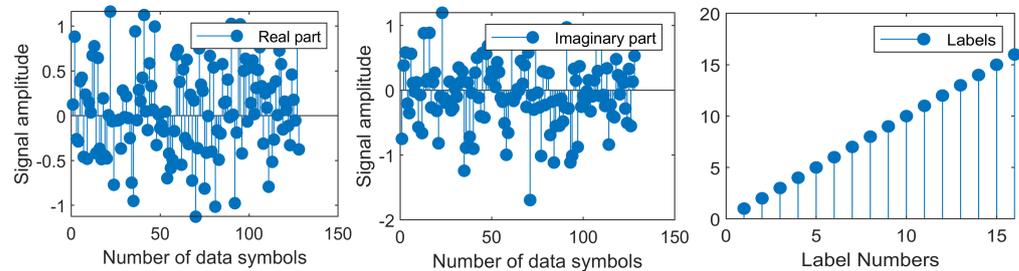


Figure 3. Dataset examples including real and imaginary parts with its labels.

2.4.2. Model Construction and Operation

The proposed model is comprised of two different networks such as 1-D CNN and the GRU network. The proposed CNN-GRU-based network design is shown in Figure 4. The input layer, convolutional layer, pooling layer, GRU layer, and output layer make up the proposed model’s structure. The input OFDM demodulated information is extracted by the convolutional layer, and afterward, the pooling layers are used to compensate for the features. Then, the GRU layer is utilized for memorizing features of the previous 1-D CNN output. In the end, the classification is performed by the softmax layer for GRU output. The operation of each layer of the proposed model is described as follows:

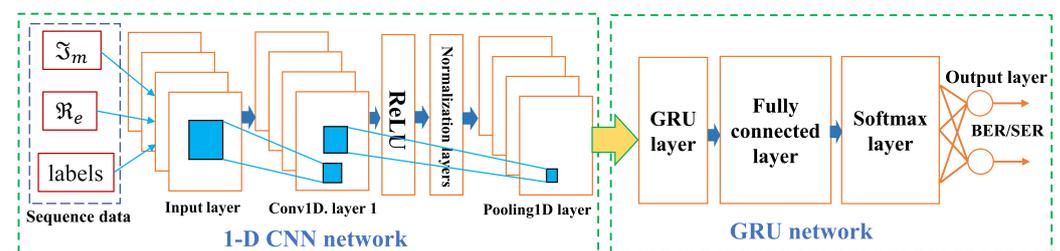


Figure 4. The architecture of the proposed CNN-GRU network: the 1-D CNN network with its layers (left side) and the GRU network layers (right side).

Input layer: We consider the input sequence features matrix of $Q = [S_1^{\Re_e}, S_1^{\Im_m}, S_2^{\Re_e}, S_2^{\Im_m}, S_3^{\Re_e}, S_3^{\Im_m}, \dots, S_N^{\Re_e}, S_N^{\Im_m}]$, which is composed with two numerical values together and the corresponding label classes. The i -th output of the sequence matrix is S_i , so that $S_0 = Q$. Eight real and imaginary values, which correspond to label 1, are contained in each sequence. The dimension of 8×1 in the input layer is assumed to equal the size of the input features. Two numerical values of sequence features with their associated labels are supplied into the CNN’s input layer from the created dataset, where the configuration sets the number of features in the input data to be equal to the input size.

Convolutional layer ($i = 1$): To extract local features from the sequence data, we utilize the convolutional layer. The convolutional layer’s parameters ($k \times s$), where k is the kernel size and s is the input data dimension, comprise a collection of learnable filters. In the convolutional layer, a total of 64 filters in 3×3 various sizes are used. The outcome feature matrix S_i may, thus, be represented as follows:

$$S_i = f(S_{i-1} \otimes w_i + b_i), \tag{12}$$

where w_i stands for the weight of the i -th layer, and b_i stands for the layer’s bias. The following layers make use of a non-linear or piecewise linear activation function called a rectified linear unit (ReLU). If the output of the ReLU function is positive, it will be directly

entered; otherwise, it will be 0. The ReLU function is shown mathematically as follows:
 $f(x) = \max(0, x)$.

Pooling layer (i = 2): The main function of the pooling layer is to compress the features that the convolutional layer has recovered in order to decrease the information dimension and the risk of network overfitting. This layer’s function in the study is to cut down on the computing time required to process sequence characteristics from earlier hidden layers using a 1-D global average pooling approach. The output of the pooling layer is the highest of the earlier features matrix. Consequently, the outcome feature matrix S_i may be expressed as follows:

$$S_i = f_p(S_{i-1}), \tag{13}$$

where the pooling function f_p is used. The dimension of S_2 is calculated from the pooling layer as $m/z \times n$, where z represents the scale value of the current layer of the pooling layer, m represents the number of filters, and n represents the number of input data time steps.

GRU layer (i = 3): The GRU layer train is the feature vectors that the CNN extracts where the GRU layer retains the internal workings of many features. According to the structure of the GRU layer, the GRU operating system may be summed up in the next following sentences. The input state information X_t at the present instant and the hidden layer information H_{t-1} that is previously learned are used to compute the update U_t and reset gate R_t in the first step. The reset gate R_t is used in the second phase to count the number of new pieces of information that are stored in the node \widehat{H}_t . The third step involves using the update gate U_t to determine the hidden layer output at the present time. The GRU computation procedure is explained using the formulas below:

$$U_t = \sigma(W_U X_t + V_U H_{t-1} + b_U) \tag{14}$$

$$R_t = \sigma(W_R X_t + V_R H_{t-1} + b_R) \tag{15}$$

$$\widehat{H}_t = \tanh(W_H X_t + V_H (R_t \otimes H_{t-1}) + b_H) \tag{16}$$

$$H_t = (1 - U_t) \otimes H_{t-1} + U_t \otimes \widehat{H}_t, \tag{17}$$

where the values for the bias are b_U , b_R , and b_H ; The gate activation function is calculated using the sigmoid function, which is denoted by $\sigma(c) = (1 + e^{-c})^{-1}$. The hyperbolic tangent function known as (tanh) is in charge of calculating the state activation function. W_U , W_R , W_H , V_U , V_R , and V_H are weight matrices. While the input state X_t and the output of the hidden layer at the previous instant are merged to generate \widehat{H}_t , the output of the hidden layer at the current instant is represented by H_t . The Hadamard product is represented as \otimes for element-wise multiplication.

Fully connected layer (i = 4): The task of activation with the softmax is provided to the fully linked layer, which plays a significant role in classification, and the final classification is performed in this layer. Here, the model calculates the probability that each sample now belongs to one of the classes and then derives the expression of a feature (Y_{pred}), which is stated as follows:

$$Y_{pred}(i) = f(L = l_i | S_3; (W, b)), \tag{18}$$

where the GRU features are S_3 and the activation function for softmax is $f(\cdot)$. The computation output for the i -th classes of the input data is represented by the variables l_i , where W and b stand for the weight and bias values, respectively.

The cross-entropy operation calculates the cross-entropy loss for single-label and multi-label classification tasks between network predictions and target values. Cross-entropy is frequently the best choice when output probability models are dealt with. Furthermore, L_2 regularization may be viewed as a successful compromise between the identification of tiny weights and the reduction of the cost function [51]. Therefore, overfitting is avoided

by using cross-entropy and $L2$ regularization. The following formulation of the loss computation serves as the basis for the goal of model training:

$$\text{Loss}(W, b) = - \sum_{i=1}^{n_s} \sum_{t=1}^c (Y^{(t)}(i) * \log(Y_{pred}^{(t)}(i))) + \frac{\lambda}{2} \sum_{i=1}^{n_s} W_i^2, \quad (19)$$

where $(Y^{(t)}(i))$ represents the prospect of the known goal, $(Y_{pred}^{(t)}(i))$ is the probability that the i -th sample belongs to the t -th class, the number of samples is n_s , the number of the class is c , and finally, the term “ λ ” defines the coefficient of the regularization of $L2$. The Adam optimizing strategy is applied to reduce the loss [52].

2.4.3. Overview of Training and Testing Process

In Figure 5, an illustration of the proposed RIS-assisted MIMO wireless communication’s training and inference process is presented. After the model is successfully trained, it is validated to check for optimal performance and inference accuracy. Figure 6 shows the flow chart for the testing procedure after training. The training procedure stops and the model for inference is stored when it reaches a model accuracy of 99.95%. The training and validation performance for the 50 epoch according to its loss and accuracy are shown in Figure 7. Figure 7 illustrates that the model performs consistently after 11 epochs, demonstrating that the parameters have learned from the data by the model. During model training, the Adam optimizer is used to analyze the data at a learning rate of 0.01. The training and inference process is shown in Algorithm 1.

Algorithm 1 The training and inference process

Training process:

- 1: Load dataset χ .
- 2: split χ into a training ζ and a validation τ set at a ratio of 80% and 20%, i.e., 4:1.
- 3: **for** e: 1 to E **do**:
- 4: Pass the processed data to build the model and configure the layers.
- 5: Configure the model parameters such as learning rate ψ , maximum epochs E_{max} , minibatch size mB , gradient threshold gT , and validation frequency f_v .
- 6: Calculate the loss function using (19).
- 7: Computing the corrective parameters and obtaining the optimal performance while updating the parameters, using the Adam optimization algorithm.
- 8: **end for**
- 9: Save model for inference step.
- 10: Output: proposed model.

Inference process:

- 11: Load the trained model.
 - 12: Initialize the parameters.
 - 13: **for** s: 1 to SNR **do**:
 - 14: **for** i: 1 to Iteration **do**:
 - 15: Generate data symbol to transmit.
 - 16: Data transmit by H_{cas} and H_d channel matrix.
 - 17: Match the label classes with a trained model to classify.
 - 18: SER and BER performance with different SNR.
 - 19: **end for**
 - 20: **end for**
 - 21: Output: SER and BER results.
-

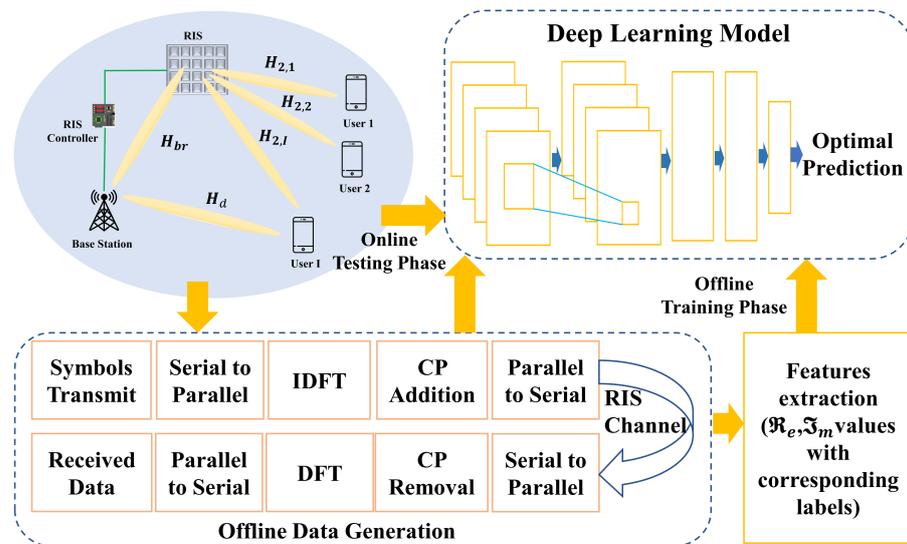


Figure 5. The structure of the proposed CNN-GRU network: training and testing phases.

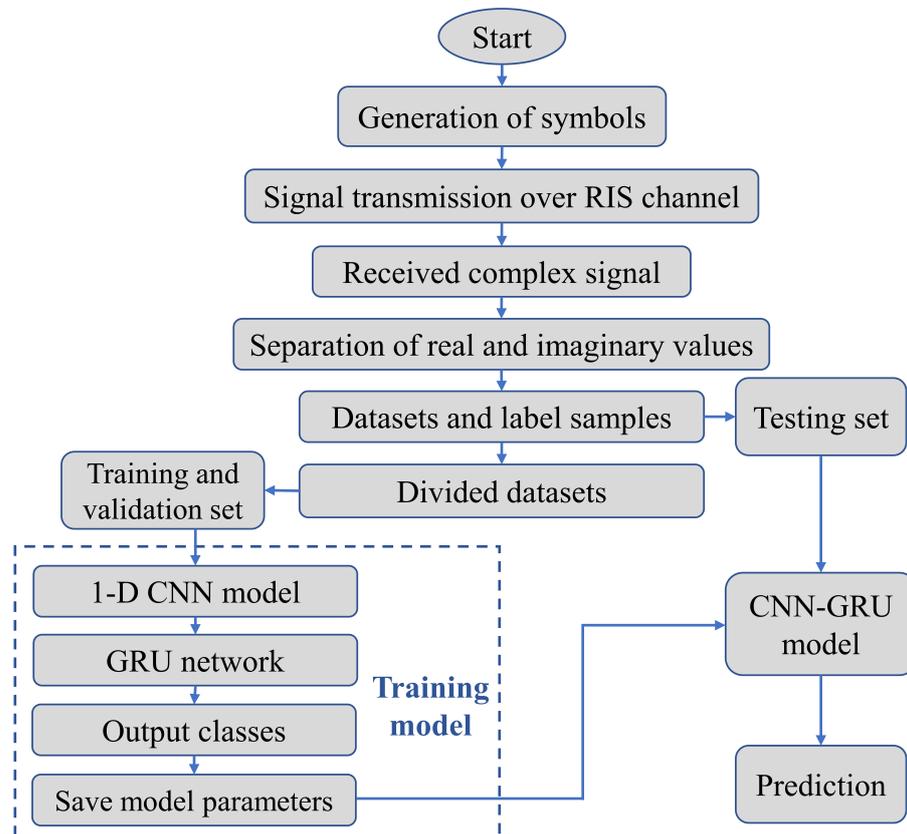


Figure 6. Flowchart for the testing process of the proposed CNN-GRU model.

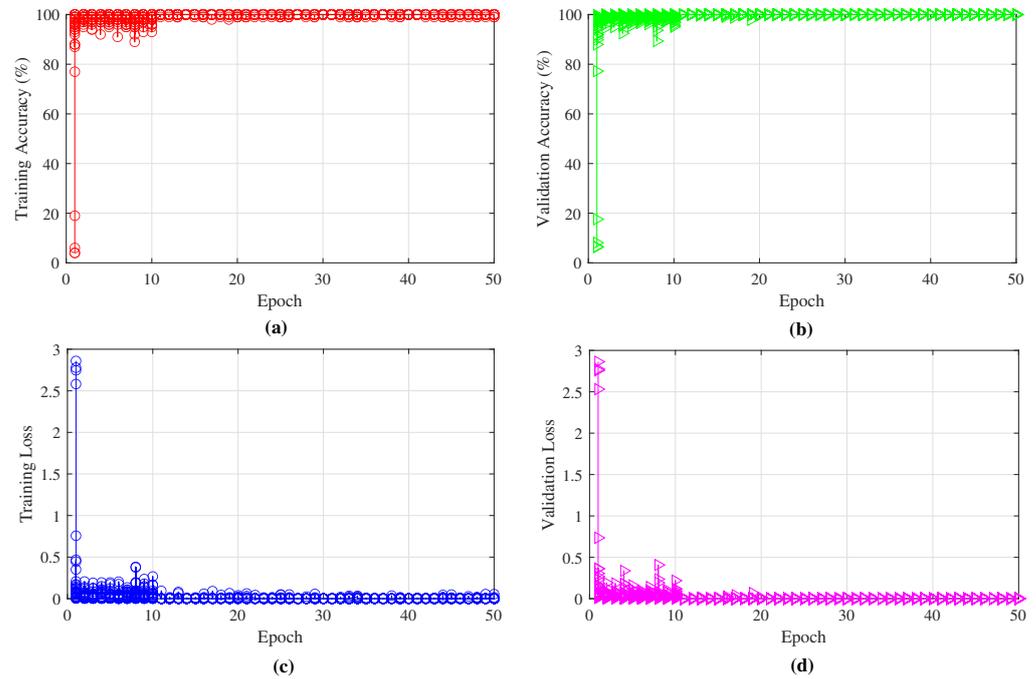


Figure 7. The model training performance analysis: (a) the training accuracy; (b) the validation accuracy; (c) the training loss; (d) the validation loss.

3. Simulation Results

The simulation results of the proposed DL-based strategy are provided in this part for wireless MIMO communication systems with RIS assistance. The simulation work is performed with MATLAB and the Windows 10 Pro operating system. The DL layers are connected to create the DNN. The DL Toolbox™ creates the DL model and keeps track of the training procedure. The training performance is improved by using an NVIDIA graphics card. In the mmWave propagation environment, it is indicated that there are $P_1 = 2$ and $P_2 = 4$ channel paths, respectively, between the BS and RIS and between RIS and each UE. The RIS to UEs channel $H_{2,i}^H$ is tuned to the non-line-of-sight (LOC), while the BS to RIS channel H_{br} is set to the LOC channel with 15 dB of Rician K -factor. The obstructions prevent access to the BS to UEs direct channel H_d . Additionally, this simulation takes into account a wireless system with RIS service in which transmit antennas are deployed in ULA configurations and RIS components are deployed in UPA configurations, with each transmit antenna and RIS element separated by a half-wavelength ($\lambda/2$). The dimension of RIS elements is configured as $N = 32 \times 16$. Table 1 presents the simulation settings for the proposed RIS communication system. In Sections 2.4.1 and 2.4.3, the training data generation and training and testing procedure are presented. The model training is implemented by setting the parameters listed in Table 2. The SNR value of 30dB is configured while the training datasets *chi* are being created. When evaluating the simulation’s performance during the online testing phase, the [0:4:20] dB SNR range is taken into consideration. To calculate the channel error rate, 128 quadrature phase-shift keying (QPSK) symbols are produced using OFDM. A length of 32 cyclic prefix (CP) as a guard interval is introduced to lessen ISI. In this study, the simulation is run to assess the BER and SER analysis in comparison to the proposed DL model and earlier studies. The data received by the i -th UEs are taken into account at a time for the BER and SER calculation. The average error rate (E_R) for the i -th UEs is represented by the BER and SER. The data from the i -th UEs are first demodulated individually, then $(E_R) = (UE1 + UE2, \dots, + UEI) / T$ is used to determine the (E_R) , where T is the number of the i -th UEs at the receiver terminal. The inaccurate demodulated bit at the receiver is referred to as the (E_R) for BER. (E_R) for

SER refers to the incorrect categorization of the received symbol. The effectiveness of the proposed DL model is compared with the LSTM model, Demod-CNN model [40], and TCN model [41], and traditional method in [40]. The effectiveness of the suggested CNN-GRU is improved by cascading with the GRU module with 1-D CNN and different parameter adjustments.

Table 1. Simulation parameters of the RIS design.

Parameters	Value
The number of RIS elements	32×16
BS to RIS paths	2
RIS to user paths	4
BS to user paths	2
Transmit antenna spacing	0.5λ
RIS elements spacing	0.5λ
Number of transmitter antenna	2
User number	2
BS antenna number	2
OFDM subcarrier number	128
Number of cyclic prefixes	32
Modulation type	QPSK
Channel noise	AWGN

Table 2. Simulation parameters of the proposed model.

Parameters	Value
Maximum epochs number	50
No. of filters in convolutional layer	64
No. of hidden units in GRU layer	100
Fully connected layer	16
Size of minibatch	100
Learning rate	0.01
Gradient threshold	1
Validation frequency	50
Used optimizer	Adam

3.1. Performance Evaluation

Figure 8 shows the prediction performance of the proposed CNN-GRU according to the predicated and tested data where the examined model is learned with the Adam optimizer with learning rate = 0.01 at the time of training. In Figure 8a, the confusion matrix is analyzed in terms of predicted and true classes for the evaluation of prediction accuracy. There are a total of 16 classes for prediction, and it is seen that the classification error of classes is very marginal. Additionally, the performance of the test sequence prediction between predicated and tested data is shown in Figure 8b. In Figure 8b, the orange and blue color-covered regions are presented with test data and predicted data, respectively. To be more clearly visible, the zoom view of the portion of the output is indicated by the blue arrow. It is evident that the miss prediction rate of class with time step is very low and the proposed model is highly efficient.

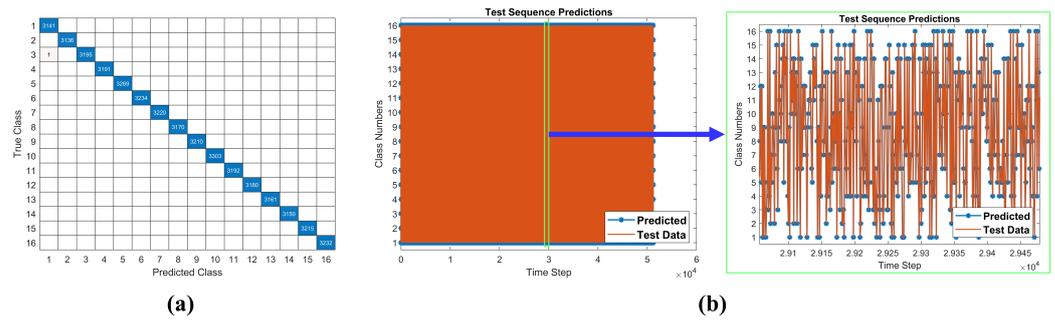


Figure 8. CNN-GRU-based RIS-assisted MIMO communication prediction performance: (a) Confusion matrix of predictions and true class; (b) Results of test sequence predictions between test and prediction data (blue arrow indicates the zoom view of a portion of the output for more visible).

Figure 9 depicts the comparison of the proposed CNN-GRU decoding performance with the traditional method, Demod-CNN model, TCN, and LSTM model according to the BER versus different SNR values. The graph absolutely exhibits that the proposed CNN-GRU-based demodulation system outperforms the traditional method, the Demod-CNN model, TCN, and LSTM model. Due to the low SNR value at the SNR (0–5) dB range, the proposed conclusions are less precise but not less than those of the other systems. However, the model performance improves as the SNR values rise up to (5–20) dB. Evidently, the proposed model significantly outperforms other systems at SNR ranging from 0 to 20 dB and BER reaches approximately 10^{-5} . In contrast, conventional and Demod-CNN systems cover SNR ranging from 0 to 30 dB and BER reaches around 10^{-4} . Up to 5 dB, the BER for the proposed CNN-GRU model follows a similar pattern to all of the methods. After this threshold, the BER for the proposed CNN-GRU model significantly decreases and ends at 20 dB SNR. Additionally, the CNN-GRU is also compared with the LSTM model and TCN model. It can be shown that the suggested model performs better than the LSTM and TCN models with the SNR (0–20) dB range. Because of the low SNR values, the LSTM model performance is degraded more than all other methods, but after 12 dB SNR, its performance follows the proposed CNN-GRU model. In addition, TCN shows a little better performance with (0–3) dB SNR, and after that, the performance is degraded and follows the LSTM trends. This demonstrates how effectively the proposed CNN-GRU model can demodulate and enhance the BER system performance.

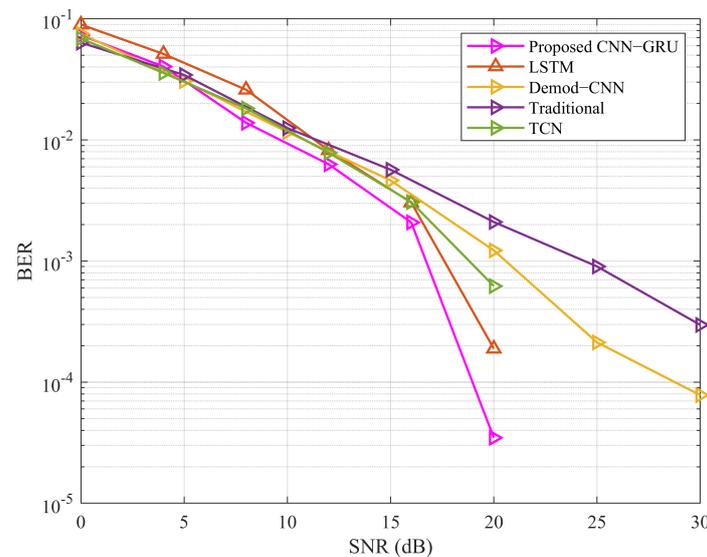


Figure 9. BER simulation results of CNN-GRU-based RIS-assisted MIMO communication.

Figure 10 shows the comparison of the proposed CNN-GRU decoding performance with the traditional demodulation system, Demod-CNN model, and LSTM model according to the SER versus different SNR values. The graph shows that the proposed CNN-GRU-based demodulation system outperforms all other methods. Evidently, the proposed model vastly outperforms other systems at SNR ranging from 0 to 20 dB and SER reaches approximately 10^{-4} . Contrarily, the conventional method and the Demod-CNN model have SNR ranges of 0 to 30 dB and SER of approximately 10^{-3} and 10^{-4} , respectively. Unlike the BER accuracy of the proposed CNN-GRU model, SER performance is also higher than the conventional, Demod-CNN, TCN, and LSTM models and shows significant improvement in decoding capability. However, due to the lower SNR, the LSTM model shows lower performance than the conventional and Demod-CNN models, but with higher SNR, it follows the proposed model performance. Additionally, the TCN model has the same performance with SNR range (0–6) dB, and after that, it follows the LSTM. It is seen from the above simulation outputs that the CNN-GRU gains a noticeable performance difference from the other four schemes. The summary is from the above simulation results; in order to enhance BER and SER for RIS-based systems, the proposed model can be an effective solution.

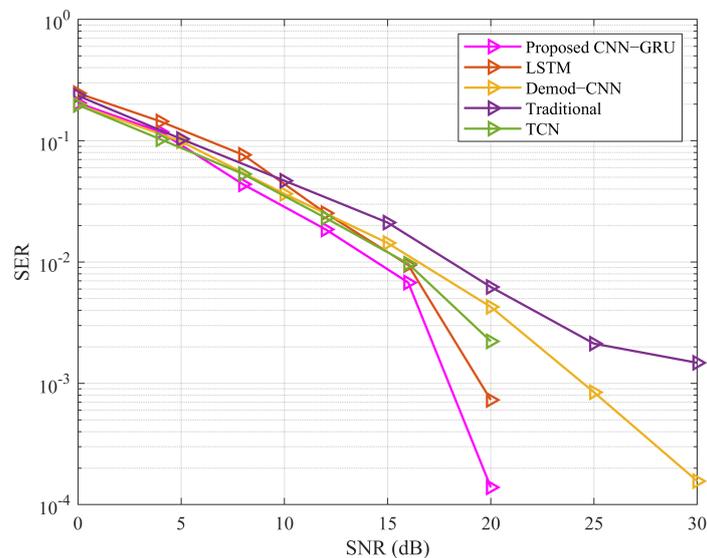


Figure 10. SER simulation results of CNN-GRU-based RIS-assisted MIMO communication.

Figure 11 shows the analysis of SER and BER of the proposed CNN-GRU model by changing different model parameters during the training of the model. The changing parameters are considered three well-known optimization algorithms, namely, Adam, SGDm, and RMSprop. In addition, to conform to the model learning rate (LR), we have configured the training parameters by considering the LR of 0.01 and 0.001, respectively. Figure 11a,b indicate that the simulation results are taken under the LR of 0.01 and three optimization algorithms. In Figure 11a, with the same LR, the Adam optimizer shows better SER performance than the others with different SNR values, and among SDGm and RMSprop, RMSprop achieves a better SER performance than SGDm. A similar performance happens for BER of Figure 11b. On the other hand, the simulation results of SER and BER with LR of 0.001 and three optimizers are shown in Figure 11c,d. Figure 11c,d indicate that with LR of 0.001, the SER and BER performance of the proposed CNN-GRU model is less than LR 0.01. Furthermore, the Adam optimizer shows a better performance compared to the others. It is indicated that in the lower SNR part, the three optimization algorithms show similar performance and follow the same trend, which confirms the proposed model robustness with three optimization algorithms.

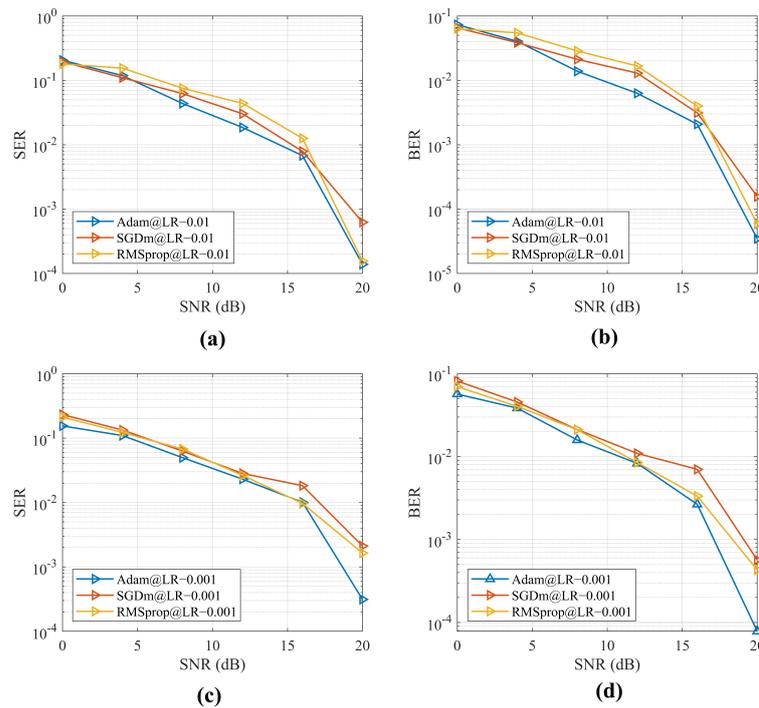


Figure 11. SER and BER simulation results of CNN-GRU model with different learning rate (LR) and optimization algorithms for RIS-assisted MIMO communication: (a) SER with LR = 0.01; (b) BER with LR = 0.01; (c) SER with LR = 0.001; (d) BER with LR = 0.001.

Figure 12 shows the throughput analysis comparison of the proposed CNN-GRU with the traditional method, Demod-CNN model, TCN, and LSTM model regarding different SNR values. The graph demonstrates that the suggested model performs almost as well as other approaches in the SNR regions of (0–5) dB. At the initial values of the SNR, the achievable rate of the CNN-GRU is 1.69 bps, which is similar to the other methods and is 1.85 bps with 5 dB SNR. After 5 dB SNR, the proposed CNN-GRU outperforms the other methods, and it achieves the maximum rate at a 20 dB SNR value. Thus, the proposed model exhibits the robustness of handling optimal prediction rates in different SNR ranges.

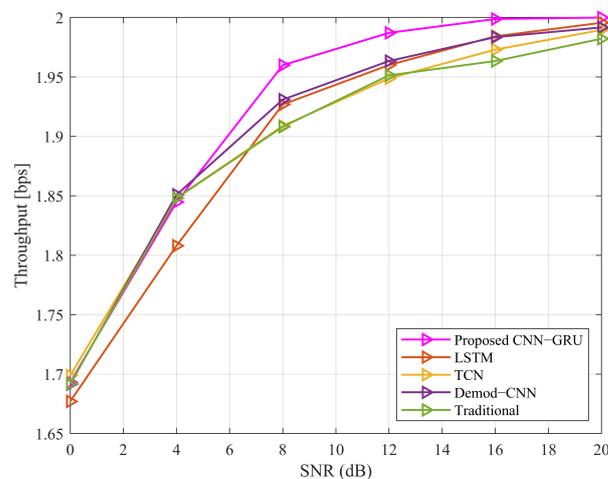


Figure 12. Throughput simulation results of CNN-GRU-based RIS-assisted MIMO communication.

3.2. Complexity Analysis

The suggested combined model's computational complexity is represented by the following formula: $\mathcal{O}(B_s \times E_s(r_c \times p_c \times t_c) + t_s t_h(3t_i + 3t_h + 3))$, where B_s represents the number of input packets obtained, E_s represents the size of an OFDM block, and r_c represents the size of the CNN input; the number of filter size and neuron size of CNN are p_c and t_c , respectively; the number of features in the input vector is t_i ; t_s is the input time sequence size; and t_h is the number of hidden units. On the other hand, the computational complexity of Demod-CNN [40] can be defined as $\mathcal{O}(B_s \times E_s(r_c \times p_c \times t_c) \times 2)$. Additionally, the computational complexity of the traditional OFDM system can be expressed as $\mathcal{O}(M_s)$, where the modulation order is M_s [53]. Because only IFFT and FFT are employed, the traditional OFDM is computationally efficient. We consider the parameter values to be $B_s = 100$, $E_s = 128$, $r_c = 8 \times 1$, $p_c = 3 \times 64$, $t_c = 100$, $t_s = 64$, $t_i = 300 \times 64$, and $t_h = 300 \times 100$ for example evaluation. After the consideration of these parameters using the Demod-CNN expression, we require 3.9×10^9 operations. In the case of the proposed model, we need 2.8×10^{11} operations for the given parameters. The research above demonstrates that the time complexity of the suggested model is greater than that of the conventional technique and the Demod-CNN model. Although the proposed model requires more complexity than others, it outperforms them all in terms of performance, and GPU parallel computing can speed up the process.

4. Conclusions

In this paper, a combined DL network to obtain the optimal decoding rate is proposed in the RIS-assisted MIMO wireless system. The proposed model is configured with a 1-D CNN model and GRU module for acquiring the maximum prediction rate. The proposed model is trained with simulated OFDM data over the RIS-assisted channel network in the offline phase. Then, the learned model is deployed in the online phase to retrieve the original data in the UE terminal to test the efficiency of the proposed model. As a result, it is possible to extract the transmitted data and determine the BER and SER performance. The error performance of the proposed model is evaluated with the well-known Adam optimization algorithm. According to the simulation results, the proposed method has a 10^{-5} BER and 10^{-4} SER in 20 dB and is better compared to other benchmarking models. In addition, the proposed CNN-GRU model has the highest throughput of 2 in 20 dB SNR. Thus, it is shown that the performance of the proposed model outperforms the conventional, Demod-CNN, TCN, and LSTM models in terms of BER, SER, and throughput. The proposed system might offer a way to make the 5G communication system better. Future applications of the proposed model could include more advanced, complicated systems for RIS-assisted wireless communication systems.

Author Contributions: Conceptualization, M.H.R. and M.A.S.S.; methodology, M.H.R.; software, M.H.R., M.A.S.S. and M.A.A.; validation, M.H.R. and M.A.S.S.; formal analysis, M.H.R., M.A.S.S. and M.A.A.; investigation, M.H.R. and M.A.S.S.; resources, H.-K.S.; data curation, M.H.R. and M.A.S.S.; writing—original draft preparation, M.H.R. and M.A.S.S.; writing—review and editing, M.H.R., M.A.S.S., M.A.A., D.-S.K. and H.-K.S.; visualization, M.H.R., M.A.S.S. and M.A.A.; supervision, Y.-H.Y., D.-S.K. and H.-K.S.; project administration, Y.-H.Y., D.-S.K. and H.-K.S.; funding acquisition, H.-K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by the ICT R&D Program of MSIT/IITP [IITP-2022-2021-0-01816, A Research on Core Technology of Autonomous Twins for Metaverse] and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1A6A1A03038540).

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict 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.* **2019**, *34*, 134–142. [[CrossRef](#)]
2. Wu, Q.; Zhang, R. Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network. *IEEE Commun. Mag.* **2019**, *58*, 106–112. [[CrossRef](#)]
3. Basar, E.; Di Renzo, M.; De Rosny, J.; Debbah, M.; Alouini, M.S.; Zhang, R. Wireless communications through reconfigurable intelligent surfaces. *IEEE Access* **2019**, *7*, 116753–116773. [[CrossRef](#)]
4. Di Renzo, M.; Zappone, A.; Debbah, M.; Alouini, M.S.; Yuen, C.; De Rosny, J.; Tretyakov, S. Smart radio environments empowered by reconfigurable intelligent surfaces: How it works, state of research, and the road ahead. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2450–2525. [[CrossRef](#)]
5. Basharat, S.; Hassan, S.A.; Pervaiz, H.; Mahmood, A.; Ding, Z.; Gidlund, M. Reconfigurable intelligent surfaces: Potentials, applications, and challenges for 6G wireless networks. *IEEE Wirel. Commun.* **2021**, *28*, 184–191. [[CrossRef](#)]
6. Shin, B.S.; Oh, J.H.; You, Y.H.; Hwang, D.D.; Song, H.K. Limited Channel Feedback Scheme for Reconfigurable Intelligent Surface Assisted MU-MIMO Wireless Communication Systems. *IEEE Access* **2022**, *10*, 50288–50297. [[CrossRef](#)]
7. Do, T.N.; Kaddoum, G.; Nguyen, T.L.; Da Costa, D.B.; Haas, Z.J. Aerial reconfigurable intelligent surface-aided wireless communication systems. In Proceedings of the 2021 IEEE 32nd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), Helsinki, Finland, 13–16 September 2021; IEEE: Piscataway Township, NJ, USA, 2021; pp. 525–530.
8. Huang, C.; Yang, Z.; Alexandropoulos, G.C.; Xiong, K.; Wei, L.; Yuen, C.; Zhang, Z.; Debbah, M. Multi-hop RIS-empowered terahertz communications: A DRL-based hybrid beamforming design. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 1663–1677. [[CrossRef](#)]
9. Jung, J.S.; Park, C.Y.; Oh, J.H.; Song, H.K. Intelligent reflecting surface for spectral efficiency maximization in the multi-user MISO communication systems. *IEEE Access* **2021**, *9*, 134695–134702. [[CrossRef](#)]
10. Guo, H.; Liang, Y.C.; Chen, J.; Larsson, E.G. Weighted sum-rate maximization for reconfigurable intelligent surface aided wireless networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 3064–3076. [[CrossRef](#)]
11. Yang, Y.; Zheng, B.; Zhang, S.; Zhang, R. Intelligent reflecting surface meets OFDM: Protocol design and rate maximization. *IEEE Trans. Commun.* **2020**, *68*, 4522–4535. [[CrossRef](#)]
12. Özdoğan, Ö.; Björnson, E.; Larsson, E.G. Intelligent reflecting surfaces: Physics, propagation, and pathloss modeling. *IEEE Wirel. Commun. Lett.* **2019**, *9*, 581–585. [[CrossRef](#)]
13. You, C.; Zheng, B.; Zhang, R. Channel estimation and passive beamforming for intelligent reflecting surface: Discrete phase shift and progressive refinement. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 2604–2620. [[CrossRef](#)]
14. Basar, E. Reconfigurable intelligent surface-based index modulation: A new beyond MIMO paradigm for 6G. *IEEE Trans. Commun.* **2020**, *68*, 3187–3196. [[CrossRef](#)]
15. Wei, L.; Huang, C.; Alexandropoulos, G.C.; Wei, E.; Zhang, Z.; Debbah, M.; Yuen, C. Multi-user holographic MIMO surfaces: Channel modeling and spectral efficiency analysis. *IEEE J. Sel. Top. Signal Process.* **2022**, *16*, 1112–1124. [[CrossRef](#)]
16. Tan, X.; Sun, Z.; Koutsonikolas, D.; Jornet, J.M. Enabling indoor mobile millimeter-wave networks based on smart reflect-arrays. In Proceedings of the IEEE INFOCOM 2018-IEEE Conference on Computer Communications, Honolulu, HI, USA, 16–19 April 2018; IEEE: Piscataway Township, NJ, USA, 2018; pp. 270–278.
17. Ye, J.; Guo, S.; Alouini, M.S. Joint reflecting and precoding designs for SER minimization in reconfigurable intelligent surfaces assisted MIMO systems. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 5561–5574. [[CrossRef](#)]
18. Huang, C.; Zappone, A.; Alexandropoulos, G.C.; Debbah, M.; Yuen, C. Reconfigurable intelligent surfaces for energy efficiency in wireless communication. *IEEE Trans. Wirel. Commun.* **2019**, *18*, 4157–4170. [[CrossRef](#)]
19. Zhang, Z.; Dai, L. A joint precoding framework for wideband reconfigurable intelligent surface-aided cell-free network. *IEEE Trans. Signal Process.* **2021**, *69*, 4085–4101. [[CrossRef](#)]
20. Jiang, T.; Cheng, H.V.; Yu, W. Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation. *IEEE J. Sel. Areas Commun.* **2021**, *39*, 1931–1945. [[CrossRef](#)]
21. Jung, M.; Saad, W.; Jang, Y.; Kong, G.; Choi, S. Performance analysis of large intelligent surfaces (LISs): Asymptotic data rate and channel hardening effects. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 2052–2065. [[CrossRef](#)]
22. Dong, P.; Zhang, H.; Li, G.Y.; Gaspar, I.S.; NaderiAlizadeh, N. Deep CNN-based channel estimation for mmWave massive MIMO systems. *IEEE J. Sel. Top. Signal Process.* **2019**, *13*, 989–1000. [[CrossRef](#)]
23. Ye, H.; Li, G.Y.; Juang, B.H. Power of deep learning for channel estimation and signal detection in OFDM systems. *IEEE Wirel. Commun. Lett.* **2017**, *7*, 114–117. [[CrossRef](#)]
24. Sejan, M.A.S.; Rahman, M.H.; Shin, B.S.; Oh, J.H.; You, Y.H.; Song, H.K. Machine Learning for Intelligent-Reflecting-Surface-Based Wireless Communication towards 6G: A Review. *Sensors* **2022**, *22*, 5405. [[CrossRef](#)] [[PubMed](#)]
25. Rahman, M.H.; Sejan, M.A.S.; Yoo, S.G.; Kim, M.A.; You, Y.H.; Song, H.K. Multi-User Joint Detection Using Bi-Directional Deep Neural Network Framework in NOMA-OFDM System. *Sensors* **2022**, *22*, 6994. [[CrossRef](#)]
26. Zappone, A.; Di Renzo, M.; Debbah, M. Wireless networks design in the era of deep learning: Model-based, AI-based, or both? *IEEE Trans. Commun.* **2019**, *67*, 7331–7376. [[CrossRef](#)]
27. Huang, C.; Mo, R.; Yuen, C. Reconfigurable intelligent surface assisted multiuser MISO systems exploiting deep reinforcement learning. *IEEE J. Sel. Areas Commun.* **2020**, *38*, 1839–1850. [[CrossRef](#)]

28. Rahman, M.H.; Sejan, M.A.S.; Aziz, M.A.; Baik, J.I.; Kim, D.S.; Song, H.K. Deep Learning Based Improved Cascaded Channel Estimation and Signal Detection for Reconfigurable Intelligent Surfaces-Assisted MU-MISO Systems. *IEEE Trans. Green Commun. Netw.* **2023**. [[CrossRef](#)]
29. Jiang, H.; Dai, L.; Hao, M.; Mackenzie, R. End-to-End Learning for RIS-Aided Communication Systems. *IEEE Trans. Veh. Technol.* **2022**, *71*, 6778–6783. [[CrossRef](#)]
30. Khan, S.; Khan, K.S.; Haider, N.; Shin, S.Y. Deep-learning-aided detection for reconfigurable intelligent surfaces. *arXiv* **2019**, arXiv:1910.09136.
31. Gao, S.; Dong, P.; Pan, Z.; Li, G.Y. Deep multi-stage CSI acquisition for reconfigurable intelligent surface aided MIMO systems. *IEEE Commun. Lett.* **2021**, *25*, 2024–2028. [[CrossRef](#)]
32. Liu, S.; Gao, Z.; Zhang, J.; Di Renzo, M.; Alouini, M.S. Deep denoising neural network assisted compressive channel estimation for mmWave intelligent reflecting surfaces. *IEEE Trans. Veh. Technol.* **2020**, *69*, 9223–9228. [[CrossRef](#)]
33. Huang, C.; Alexandropoulos, G.C.; Yuen, C.; Debbah, M. Indoor signal focusing with deep learning designed reconfigurable intelligent surfaces. In Proceedings of the 2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, 2–5 July 2019; IEEE: Piscataway Township, NJ, USA, 2019; pp. 1–5.
34. Taha, A.; Zhang, Y.; Mismar, F.B.; Alkhateeb, A. Deep reinforcement learning for intelligent reflecting surfaces: Towards standalone operation. In Proceedings of the 2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Atlanta, GA, USA, 26–29 May 2020; IEEE: Piscataway Township, NJ, USA, 2020; pp. 1–5.
35. Liao, G.P.; Gao, W.; Yang, G.J.; Guo, M.F. Hydroelectric generating unit fault diagnosis using 1-D convolutional neural network and gated recurrent unit in small hydro. *IEEE Sens. J.* **2019**, *19*, 9352–9363. [[CrossRef](#)]
36. Yang, B.; Cao, X.; Huang, C.; Yuen, C.; Qian, L.; Di Renzo, M. Intelligent spectrum learning for wireless networks with reconfigurable intelligent surfaces. *IEEE Trans. Veh. Technol.* **2021**, *70*, 3920–3925. [[CrossRef](#)]
37. Xu, M.; Zhang, S.; Zhong, C.; Ma, J.; Dobre, O.A. Ordinary differential equation-based CNN for channel extrapolation over RIS-assisted communication. *IEEE Commun. Lett.* **2021**, *25*, 1921–1925. [[CrossRef](#)]
38. Zhang, S.; Zhang, S.; Gao, F.; Ma, J.; Dobre, O.A. Deep learning optimized sparse antenna activation for reconfigurable intelligent surface assisted communication. *IEEE Trans. Commun.* **2021**, *69*, 6691–6705. [[CrossRef](#)]
39. Gupta, K.D.; Nigam, R.; Sharma, D.K.; Dhurandher, S.K. LSTM-Based Energy-Efficient Wireless Communication With Reconfigurable Intelligent Surfaces. *IEEE Trans. Green Commun. Netw.* **2022**, *6*, 704–712. [[CrossRef](#)]
40. Sejan, M.A.S.; Rahman, M.H.; Song, H.K. Demod-CNN: A Robust Deep Learning Approach for Intelligent Reflecting Surface-Assisted Multiuser MIMO Communication. *Sensors* **2022**, *22*, 5971. [[CrossRef](#)]
41. Sejan, M.A.S.; Rahman, M.H.; Aziz, M.A.; You, Y.H.; Song, H.K. Temporal Neural Network Framework Adaptation in Reconfigurable Intelligent Surface-Assisted Wireless Communication. *Sensors* **2023**, *23*, 2777. [[CrossRef](#)]
42. Zhao, R.; Wang, D.; Yan, R.; Mao, K.; Shen, F.; Wang, J. Machine health monitoring using local feature-based gated recurrent unit networks. *IEEE Trans. Ind. Electron.* **2017**, *65*, 1539–1548. [[CrossRef](#)]
43. Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation. *arXiv* **2014**, arXiv:1406.1078.
44. Helmy, I.; Tarafder, P.; Choi, W. LSTM-GRU Model-Based Channel Prediction for One-Bit Massive MIMO System. *IEEE Trans. Veh. Technol.* **2023**, 1–6. [[CrossRef](#)]
45. Ali, M.H.E.; Rabeh, M.L.; Hekal, S.; Abbas, A.N. Deep Learning Gated Recurrent Neural Network-Based Channel State Estimator for OFDM Wireless Communication Systems. *IEEE Access* **2022**, *10*, 69312–69322.
46. Chen, J.; Liang, Y.C.; Cheng, H.V.; Yu, W. Channel estimation for reconfigurable intelligent surface aided multi-user MIMO systems. *arXiv* **2019**, arXiv:1912.03619.
47. Taha, A.; Alrabeiah, M.; Alkhateeb, A. Enabling large intelligent surfaces with compressive sensing and deep learning. *IEEE Access* **2021**, *9*, 44304–44321. [[CrossRef](#)]
48. Busari, S.A.; Huq, K.M.S.; Mumtaz, S.; Dai, L.; Rodriguez, J. Millimeter-wave massive MIMO communication for future wireless systems: A survey. *IEEE Commun. Surv. Tutorials* **2017**, *20*, 836–869. [[CrossRef](#)]
49. Wei, X.; Shen, D.; Dai, L. Channel estimation for RIS assisted wireless communications—Part II: An improved solution based on double-structured sparsity. *IEEE Commun. Lett.* **2021**, *25*, 1403–1407. [[CrossRef](#)]
50. Björnson, E.; Larsson, E.G.; Marzetta, T.L. Massive MIMO: Ten myths and one critical question. *IEEE Commun. Mag.* **2016**, *54*, 114–123. [[CrossRef](#)]
51. Van Laarhoven, T. L2 regularization versus batch and weight normalization. *arXiv* **2017**, arXiv:1706.05350.
52. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.
53. Jaradat, A.M.; Hamamreh, J.M.; Arslan, H. Modulation options for OFDM-based waveforms: Classification, comparison, and future directions. *IEEE Access* **2019**, *7*, 17263–17278. [[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.