Demod-CNN: A Robust Deep Learning Approach for Intelligent Reflecting Surface-Assisted Multiuser MIMO Communication

Mohammad Abrar Shakil Sejan 1,2,†, Md Habibur Rahman 1,2,† and Hyoung-Kyu Song 1,2,*

1 Department of Information and Communication Engineering, Sejong University, Seoul 05006, Korea
2 Department of Convergence Engineering for Intelligent Drone, Sejong University, Seoul 05006, Korea
* Correspondence: songhk@sejong.ac.kr
† These authors contributed equally to this work.

Abstract: The intelligent reflecting surface (IRS) is a novel and innovative communication technology that aims at the control of the wireless environment. The IRS is considered as a promising technology for sixth-generation wireless communication. In the last few years, machine learning has emerged as a powerful tool for solving complex problems in diverse application areas. In this paper, we propose a convolutional neural network (CNN)-based demodulation technique called Demod-CNN in IRS-based wireless communication for multiple users. A multiple-input multiple-output based orthogonal multiple frequency division multiplexing system is considered for channel modeling. The received signal data are used for training and testing the model. The simulation results show that the proposed model performs better than the conventional demodulation technique.

Keywords: intelligent reflecting surface; demodulation; convolutional neural network; MIMO; OFDM

1. Introduction

The intelligent reflecting surface (IRS) is considered as a novel smart radio technology for wireless communication systems [1]. The IRS is a two-dimensional metasurface of electromagnetic material that is composed of a large array of passive scattering elements with a specially designed physical structure [2]. Each element in the IRS can be controlled in a software-defined manner to change the electromagnetic proprieties (e.g., phase shift, amplitude) of the reflecting RF signal. This provides a unique advantage for non-line-of-sight communication to achieve a strong signal reception using the IRS [3]. One of the important properties of the IRS is that it can be implemented with low hardware cost and energy consumption for high gain beamforming [4–6]. Thus, it is desired to use the IRS in implementing millimeter wave communication for the sixth generation (6G). In addition, the IRS can be a potential candidate to enhance wireless sensor network (WSN) communication as it has the following advantages [7]: the IRS is composed of a nearly passive device electromagnetic material that is low cost, the IRS can reconfigure the wireless propagation environment to work as an alternative link when a direct link is not available, the IRS is energy-efficient compared to traditional relaying schemes, and the IRS can be implemented as full-duplex communication. Each of these features is desirable for WSN systems.

Multiple-input-multiple-output (MIMO)-based communication has been widely adopted in wireless communication [8,9]. MIMO-based communication has also been introduced to IRS-based communication [10,11]. Machine learning (ML)-based methods have also gained significant attention from researchers [12,13]. ML techniques have become very popular in solving complex problems for enhancing wireless communication systems without explicit programming [14]. A part of ML, deep learning (DL)-based schemes are also effective in predicting the channel estimation, beamforming optimization, and conducting phase
optimization for IRS-based communication. The convolutional neural network (CNN) is a well-known DL architecture that can provide effective solutions for different problem domains. In [15], authors proposed a denoising-convolution-neural-network-based channel estimation model for MIMO IRS. The transmitted complex data are received as a noisy image to increase the channel accuracy. The study in [16] proposed a twin CNN for the estimation of the direct channel and the cascade channel for IRS. The CNN network is trained with pilot data and can tolerate a user location of about 4 degrees. Another study in [17] used a CNN-based approach to estimate millimeter wave channel estimation. In addition to channel estimation, CNN can be applied in performance analysis such as bit error rate (BER) or symbol error rate (SER) for IRS-based communication.

Motivated from above, in this paper we propose a CNN-based ML architecture to improve the BER and SER performance of the MIMO-based IRS system. Orthogonal frequency division multiplexing (OFDM) based pilot symbols are transmitted through a frequency selective fading channel. Both IRS-based cascade channel model and direct channel are considered for communication. A CNN-based demodulation technique was proposed named Demod-CNN (Demodulation using CNN) to improve the channel performance in the case of MIMO communication. The proposed CNN model is trained with received complex data to map into bits for the channel in different signal-to-noise ratio (SNR) values. The trained model was tested with different datasets for performance evaluation.

The contributions of this study can be listed as:

• An IRS-based MIMO channel configuration system is considered for wireless communication to test machine learning assisted demodulation.
• A CNN-based demodulation technique Demod-CNN is proposed to demodulate the received signal.

Notations: The lower-case and upper-case boldface letter $h$ and $H$ represent a vector and matrix, respectively; $H^H$ denotes the conjugate transpose matrix of $H$; $\text{diag}(x)$ denotes the diagonal matrix having vector $x$ on its diagonal; and $\otimes$ is the Kronecker product.

2. System Model

2.1. OFDM Communication

OFDM is a multi-carrier modulation technique applied in wireless communication for high spectral efficiency and good performance against frequency selective fading. The modulation process for OFDM starts by converting the binary inputs into phase shift Keying (PSK) or quadrature amplitude modulation for the purpose of mapping $D$ parallel streams. $O$ is the number of subcarriers in the OFDM system. If the value of $O$ is increased, the performance of the model is degraded. Thus, the value of $O$ should be chosen optimally depending on the required system. Let $X_i[q]$ be the $i$-th transmit symbol for the $q$-th subcarrier, where $i = 0, 1, 2, \ldots, \infty$ and $q = 0, 1, 2, \ldots, O - 1$. Next, the symbols are converted to the frequency domain from the time domain by using inverse fast fourier transform (IFFT). To avoid inter symbol interference, a cyclic prefix is added in the signal. One single symbol duration is considered as $T_s$, so for $Q$, symbol transmission requires $T_{sym} = QT_s$. Thus, for the $i$-th symbol and the $q$-th subcarrier, the OFDM signal $Y_{i,q}(t)$ is written as follows:

$$Y_{i,q}(t) = \begin{cases} e^{j2\pi f_q(t-T_{sym})}, & 0 < t \leq T_{sym} \\ 0, & \text{elsewhere} \end{cases}$$

where $f_q$ is the center frequency for the $q$-th subcarrier. The continuous time domain signal can be expressed as follows [18]:

$$X_i(t) = \sum_{i=0}^{\infty} \sum_{q=0}^{O-1} X_i[q] e^{j2\pi f_q(t-T_{sym})}.$$ (2)

The transmitted symbol can be considered as follows:
where $f_q = q / T_{sym}$.

### 2.2. IRS Based Communication

In this section, we describe the IRS network architecture and other related considerations. Figure 1 shows a communication scenario with the IRS where the base station (BS) has $M$ uniform planar array (UPA) antenna and each user (UE) has a single antenna. The number of IRS elements is considered as $N$. The received signal via IRS for each UE can be expressed as follows [19]:

$$y_i = H_{ui}^H \Psi H_b x + n, \quad (4)$$

where $y_i$ is the received signal at the $i$-th UE, the transmitted signal is $x \in \mathbb{C}^{M \times 1}$, $H_b \in \mathbb{C}^{N \times M}$ is the channel matrix between BS and IRS, $H_{ui} \in \mathbb{C}^{1 \times N}$ is the channel matrix from IRS to user $i$, and $n \sim \mathcal{CN}(0, \sigma^2)$ is the additive white Gaussian noise at the $i$-th user. $\Psi$ is the diagonal matrix where $\Psi = \text{diag}(\zeta) \in \mathbb{C}^{N \times N}$ represents the phase shift values of IRS elements. Each element can be defined as $\zeta = [\eta_1 e^{j\omega_1}, \eta_2 e^{j\omega_2}, \ldots, \eta_N e^{j\omega_N}] \in \mathbb{C}^{N \times 1}$, where $\eta_n \in [0, 1]$ denotes the amplitude and $\omega_n \in [0, 2\pi]$ is the phase shift coefficient of the $n$-th reflective element. Thus, $\Psi$ can be written as:

$$\Psi = \begin{bmatrix} \eta_1 e^{j\omega_1} & \ldots & \ldots & \ldots \\ \ldots & \eta_2 e^{j\omega_2} & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \eta_N e^{j\omega_N} \end{bmatrix}. \quad (5)$$

![Figure 1. Intelligent reflecting surface in the multiuser MIMO communication system.](image)

For easy calculation, the constant amplitude coefficient is considered as $\omega_n = 1$. The total channel with direct communication link can be represented by [20]:

$$y_{ii} = H_{ui}^H \Psi H_b x + H_d x + n, \quad (6)$$
where \( H_f \in \mathbb{C}^{M \times 1} \) is the direct channel between BS and UE. The channel matrices \( H_{ul}, H_{br}, \) and \( H_f \) follow the Rayleigh fading distribution, and each column \( k \) is modeled as:

\[
\begin{align*}
    h_{ul}(1,k) &= \sqrt{h_{ul}(1,k)} \quad (7a) \\
    h_{br}(1,k) &= \sqrt{h_{br}(1,k)} \quad (7b) \\
    h_{d}(1,k) &= \sqrt{h_{d}(1,k)} \quad (7c)
\end{align*}
\]

where \( \kappa \) is the path loss factor and \( \sqrt{h_{ul}(1,k)}, \sqrt{h_{br}(1,k)}, \) and \( \sqrt{h_{d}(1,k)} \) are \( \sim \mathcal{CN}(0, \sigma^2) \). In addition, we consider the Saleh–Valenzuela channel model [8], which is applicable for a multipath propagation environment. The general theory of the Saleh–Valenzuela model for mmWave communication can be modeled as follows:

\[
    h = \sqrt{\frac{N}{L}} \sum_{l=0}^{L} a_l \gamma_l^H, \quad (8)
\]

where \( h \) is the channel vector, \( a_l \) is the complex gain of the \( l \)-th path, \( L \) is the total number of paths, \( \gamma_l^H \) is the azimuth angle of departure, and \( \phi_l^H \) is the elevation angle of departure and \( a(\gamma_l^H, \phi_l^H) \) is the array response vector. For a typical \( N_1 \times N_2 \) UPA, the array response vector written as follows [21]:

\[
    a(\gamma, \phi) = \frac{1}{\sqrt{N}} [e^{-j2\pi d \sin(\gamma)} \cos(\phi) n_1 / \lambda] \otimes [e^{-j2\pi d \sin(\phi) n_2 / \lambda}], \quad (9)
\]

where \( n_1 = [0, 1, \ldots, N_1 - 1] \) and \( n_2 = [0, 1, \ldots, N_2 - 1] \), \( \lambda \) is the carrier wavelength, and \( d \) is the antenna spacing fulfilling the condition \( d = \lambda / 2 \). The BS-IRS channel \( H_b \) can be expressed as follows:

\[
    H_b = \sqrt{\frac{MN}{L_1}} \sum_{l_1=1}^{L_1} \beta_{l_1} b(\gamma_{l_1}^H, \phi_{l_1}^H) a^H(\gamma_{l_1}^H, \phi_{l_1}^H), \quad (10)
\]

where \( L_1 \) represents the number of paths between the BS and the \( i \)-th UE, \( \beta_{l_1} \) represents the complex gain of the path, \( b(\gamma_{l_1}^H, \phi_{l_1}^H) \) represents the steering vector related to the IRS, \( a(\gamma_{l_1}^H, \phi_{l_1}^H) \) is the steering vector related to BS for the \( l \)-th path.

Next, the channel between the IRS and UE can be defined as follows:

\[
    H_{ul}^H = \sqrt{\frac{N}{L_2}} \sum_{l_2=1}^{L_2} \beta_{l_2} a^H(\gamma_{l_2}^H, \phi_{l_2}^H), \quad (11)
\]

where \( L_2 \) is the number of paths between the IRS and the \( i \)-th user, \( \beta_{l_2} \) is the complex gain of paths, \( \gamma_{l_2}^H \) and \( \phi_{l_2}^H \) are the azimuth and elevation angle of departure of the signal, and \( a(\gamma_{l_2}^H, \phi_{l_2}^H) \) is the steering vector. The cascade channel for BS to UE is as follows:

\[
    H_{ca} = \sqrt{\frac{MN}{L_1 L_2}} \sum_{l_1=1}^{L_1} \sum_{l_2=1}^{L_2} \beta_{l_1} \beta_{l_2} \text{diag}(a^H(\gamma_{l_2}^H, \phi_{l_2}^H)) b(\gamma_{l_1}^H, \phi_{l_1}^H) a^H(\gamma_{l_1}^H, \phi_{l_1}^H). \quad (12)
\]

The channel matrix \( H_{ul}^H \Psi H_b \) can be expressed as follows:

\[
    H_{ul}^H \Psi H_b = H_{ul}^H \text{diag}(\xi) H_b = \xi^T \text{diag}(H_{ul}^H) H_b. \quad (13)
\]

Then, the following equation is obtained:

\[
    H_{ca} \xi^T = H_{ul}^H \Psi H_b. \quad (14)
\]

The total received signal \( y_{it} \) can be written as follows:

\[
    y_{it} = (H_{ca} \xi^T + H_d) x + n. \quad (15)
\]
2.3. Deep Learning Model

Figure 2 shows the architectural design of the proposed CNN network. For successive training with the generated dataset, in this study, we have considered 1-D CNN model architecture to evaluate BER and SER for multiuser MIMO signals. The complex signal was first separated into real and imaginary parts. Then, the two numerical values along with the label are fed to the CNN model. In the proposed model, the input layer is fed into the OFDM data symbol, where the input size is equal to the number of features of input data. The size of the input features is considered as $2 \times 2 \times 2 = 8$. The input size depends upon the number of users and the number of total antennas used in the system. The next two layers are connected by a convolutional 1D layer, ReLU activation function, and normalization layers. The first convolutional layer consists of a $3 \times 3$ filter size and total 32 numbers of filters are used. In the second convolutional layer, a $3 \times 3$ filter size, and 64 filters are used. The convolutional layer comprises a rectangular grid of neurons, where every neuron receives inputs from a rectangular part of the earlier layer. To reduce the output of the convolutional layers to a single vector, a global average pooling 1D layer is used. In the fully connected layer, each neuron is connected to all neurons in the previous layer and gathers all the features and internal information combined by the prior layers. In every time step of the CNN model, the fully connected layer works individually. We utilize the softmax activation function to derive the outputs for the final layer. In the terminal layer, we use the classification layer to map the output to vector probabilities and specify a fully connected layer with an output size matching the number of classes. The convolution operation for 1D forward propagation can be expressed as follows: [22]:

$$
\tilde{\zeta}_v^k = b_v^k + \sum_{p=1}^{N_v-1} \text{conv1D}(w_v^{v-1} \cdot s_v^{v-1}),
$$

where the input is $\zeta_v^k$, the bias of the $k$-th neuron at layer $v$ represents $b_v^k$, the output of the $p$-th neuron at layer $(v-1)$ is defined as $s_v^{v-1}$, and the kernel from the $p$-th neuron at layer $(v-1)$ to the $k$-th neuron at layer $v$ presents $w_v^{v-1}$. Hence, the dimension of the output arrays $s_v^{v-1}$ is greater than the dimension of the input array $\zeta_v^k$.

To normalize the $b$-th $\zeta_v^k$, we use the batch normalization layer [23] with the batch size $B$, which can be expressed as follows:

$$
\bar{\zeta}_v^k = \frac{1}{P V B} \sum_{p=1}^{P V} \sum_{v=1}^{V} \sum_{b=1}^{B} (\zeta_v^k - \mu_k),
$$

(17)

$$
\sigma_k^2 = \frac{1}{P V B} \sum_{p=1}^{P V} \sum_{v=1}^{V} \sum_{b=1}^{B} ((\zeta_v^k - \mu_k)^2),
$$

(18)

$$
\hat{\zeta}_v^k = \frac{\zeta_v^k - \mu}{\sqrt{\sigma^2 + \epsilon}},
$$

(19)
\[ \zeta_k^v = a\zeta_k^v + \varrho, \] (20)

where \( \mu \) and \( \sigma^2 \) present the mean and variance of \( \zeta_k^v \) respectively. \( P \) and \( V \) denote the size of the tensors for calculating the mean and variance. \( \epsilon \) is a small constant, which is negligible. In the training process, \( a \) and \( \varrho \) are the learnable parameters, which will be updated. It can be noted that the learnable parameters \( a \) and \( \varrho \) denote the whole weights in convolutional layers, batch normalization layer, and fully-connected layer, respectively.

The activation function considered as the Relu function can be written as follows:

\[
\begin{align*}
f(x) &= \begin{cases} 
0 & \text{for } x < 0, \\
x & \text{for } x \geq 0.
\end{cases}
\end{align*}
\] (21)

To activate the output signal \( \zeta_k^v \), herein, we utilize the Relu activation function, which can be formulated as follows:

\[ \zeta_k^v R = f(\zeta_k^v R). \] (22)

In addition, to determine the possibility \( \zeta_k \) of each category, we use the Softmax layer which can be expressed as follows:

\[
\text{Softmax}(\zeta_k) = \frac{\exp(\zeta_{fc k})}{\sum_{z=1}^{Z} \exp(\zeta_{fc z})}.
\] (23)

Finally, the mean-squared error (MSE) defines the loss function for the overall network, which is presented as follows:

\[
\text{MSE} = \sum_{i=1}^{N_\varphi} (\zeta_k - y_{it})^2,
\] (24)

where \( N_\varphi \) represents the number of class labels. Figure 3 shows the overall flowchart for the proposed system.

![Image](image.png)

**Figure 3.** Flowchart for the proposed system workflow.
The computational complexity of the proposed model can be specified as $O(P_s \times F_s(i_c \times f_c \times n_c) \times 2)$, where $P_s$ is the number of received input packets, $F_s$ is the OFDM block size, $i_c$ is the input size of CNN, $f_c$ is the filter size of CNN, and $n_c$ is the neuron size of CNN. In contrast, the conventional OFDM system has the complexity $O(M_s)$, where $M_s$ is the modulation order [24]. The conventional OFDM is computationally efficient because it only uses IFFT and FFT. The complexity of the proposed model is higher than the conventional OFDM, but the performance is improved.

3. Simulation Setup

For the simulation, it is assumed that the number of BS antennas is $M = 2$ and the number of IRS elements is $N = 512$. It is assumed that the number of horizontal and vertical IRS elements are 32 and 16. Two users receive the data from the BS to UE = 2. The number of paths between BS and IRS is two, and the IRS to the $i$-th UE is four. For direct BS to UE channel, two paths are considered. In this simulation, multiuser interference is not considered, and it is assumed that each UE has same channel properties. Then, the BS can demodulate different UE signals successfully. We first determine the BER using the conventional technique following the same channel configuration for IRS-based communication. The simulation parameters are listed in Table 1. Additive white Gaussian noise is added to the transmission to simulate the channel noise environment. For the downlink channel, the three-channel matrix is considered: BS-IRS, IRS-UE, and BS-UE. The total $10^6$ packets are sent each time. One hundred and twenty-eight quadrature phase-shift keying (QPSK) symbols are generated using OFDM to estimate the channel error rate. To reduce the inter-symbol interference, a length of 32 cyclic prefix is added. For BER calculation, the data received by both users are considered simultaneously.

Table 1. Simulation parameters.

| Parameters        | Value     |
|-------------------|-----------|
| IRS elements      | $32 \times 16$ |
| Transmitting antenna | 2         |
| Number of user    | 2         |
| Number of subcarrier | 128        |
| Modulation        | QPSK      |
| Number of epoch   | 100       |
| Minibatch size    | 200       |
| Input size        | 8         |
| Learning rate     | 0.01      |
| Optimizer         | ADAM      |
| Noise             | AWGN      |

The CNN-based model is trained by using the label and data generated by the QPSK symbol. The labeling is performed by combining the data from two transmitting antennas. As each antenna can generate 4 unique symbols, thus $4 \times 4 = 16$ labels are created for two data stream combinations. At the receiver, the received symbols are separated into real and imaginary parts along with the label. During the training, we capture data at the SNR 25 dB level to generate an offline input training set. We use 100,000 data sets for model development, of which $(4/5)$ are used for training and $(1/5)$ is used for validation. At 99.91% model accuracy, we stop the training process and save the model for inference. The training of the model less than 80,000 data sets does not produce the optimal performance. To achieve the highest accuracy, the model needs to have 80,000 data sets. However, 50,000 data sets can be employed for training, and in this case, the accuracy of the model is reduced. It is obvious that training with more data sets will require more time for training. Figure 4 shows the training and validation progress of the proposed model for 50 epochs. The upper left image shows the training accuracy, while the upper right image shows the validation accuracy. An aging lower left curve shows the training loss, and the lower right images shows the validation loss. It can be seen from the Figure 4 that after 20 epochs the
model shows a stable performance, which indicates the model has learned the parameters from the data. If the number of antennas, users, IRS elements, OFDM parameters, or multi-paths is changed, the retraining of the model is required. In the future, we will try to build a model that can adopt these changes with minimum training.

Figure 4. Proposed Demod-CNN model training and validation progress for 50 epochs.

4. Results and Discussion

In this section, the simulation results of the proposed model in terms of BER and SER are discussed. The BER and SER represent the average error rate for the two UEs. Both UEs data are first demodulated separately, and the error rate is calculated by error rate = (user 1 + user 2)/2. For BER, error rate refers to the wrong demodulated bit at the receiver. For SER, error rate refers to the wrong classification of the received symbol. In Figure 5, the BER curve is compared with the CNN-based demodulation against the conventional demodulation system. It is evident from the figure that CNN-based demodulation performs better than the traditional one. At SNR 0 dB, the model provides less accurate results than the conventional system because of the low SNR value. However, as the SNR values increase, the model performance improves. From SNR 5 dB, the model outperforms the conventional demodulation system. The BER for CNN-based demodulation has a similar trend to the conventional system up to 12 dB; after this point, BER improves dramatically for CNN-based demodulation. This indicates the efficient demodulation capability of the proposed CNN model. In Figure 6, SER versus SNR is plotted for the proposed model. The model shows better SER compared to the conventional demodulation scheme. After the SER 15 dB, the difference becomes clearer toward the higher SNR. Thus, the proposed CNN model can be used for improving BER and SER for IRS-based wireless communication.
5. Conclusions

In this paper, we have proposed a CNN-based model “Demod-CNN” OFDM demodulation system for IRS-aided multiuser MIMO communication. In the proposed model, we designed the input layer to receive the OFDM signal by separating real and imaginary parts of the symbol. It is then classified to recognize the transmitted bits by the classification layer. The simulation results provide that the proposed Demod-CNN can outperform the conventional demodulation system. In the future, we want to expand our model for multi-IRS systems to achieve massive MIMO communication.
Author Contributions: Conceptualization, M.A.S.S. and M.H.R.; methodology, M.A.S.S.; software, M.A.S.S. and M.H.R.; validation, M.A.S.S. and M.H.R.; formal analysis, M.A.S.S. and M.H.R.; investigation, M.A.S.S. and M.H.R.; resources, H.-K.S.; data curation, M.A.S.S. and M.H.R.; writing—original draft preparation, M.A.S.S. and M.H.R.; writing—review and editing, M.A.S.S., M.H.R., and H.-K.S.; visualization, M.A.S.S. and M.H.R.; supervision, H.-K.S.; project administration, H.-K.S.; and funding acquisition, H.-K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2022-2018-0-01423), supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation) and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2020R1A6A1A03038540).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

1. 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]
2. Gong, S.; Lu, X.; Hoang, D.T.; Niyato, D.; Shu, L.; Kim, D.I.; Liang, Y.C. Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey. *IEEE Commun. Surv. Tutor.* 2020, 22, 2283–2314. [CrossRef]
3. Wu, Q.; Zhang, S.; Zheng, B.; You, C.; Zhang, R. Intelligent reflecting surface-aided wireless communications: A tutorial. *IEEE Trans. Commun.* 2021, 69, 3313–3351. [CrossRef]
4. Wei, X.; Shen, D.; Dai, L. Channel estimation for RIS assisted wireless communications—Part I: Fundamentals, solutions, and future opportunities. *IEEE Commun. Lett.* 2021, 25, 1398–1402. [CrossRef]
5. 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]
6. Lin, Z.; Niu, H.; An, K.; Wang, Y.; Zheng, G.; Chatzinotas, S.; Hu, Y. Refracting RIS Aided Hybrid Satellite-Terrestrial Relay Networks: Joint Beamforming Design and Optimization. *IEEE Trans. Aerosp. Electron. Syst.* 2022, 58, 3717–3724. [CrossRef]
7. Liu, Y.; Liu, X.; Mu, X.; Hou, T.; Xu, J.; Di Renzo, M.; Al-Dhahir, N. Reconfigurable Intelligent Surfaces: Principles and Opportunities. *IEEE Commun. Surv. Tutor.* 2021, 23, 1546–1577. [CrossRef]
8. 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. Tutor.* 2017, 20, 836–869. [CrossRef]
9. An, K.; Lin, M.; Ouyang, J.; Zhu, W.; Member, S. Satellite Terrestrial Networks. *IEEE J. Sel. Areas Commun.* 2016, 34, 3025–3037. [CrossRef]
10. Zhang, J.; Liu, J.; Ma, S.; Wen, C.K.; Jin, S. Large System Achievable Rate Analysis of RIS-Assisted MIMO Wireless Communication with Statistical CSIT. *IEEE Trans. Wirel. Commun.* 2021, 20, 5572–5585. [CrossRef]
11. Zhang, S.; Zhang, R. Capacity Characterization for Intelligent Reflecting Surface Sided MIMO Communication. *IEEE J. Sel. Areas Commun.* 2020, 38, 1823–1838. [CrossRef]
12. 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]
13. Zhang, J.; Su, Q.; Tang, B.; Wang, C.; Li, Y. DPSNet: Multitask Learning Using Geometry Reasoning for Scene Depth and Semantics. *IEEE Trans. Neural Netw. Learn. Syst.* 2021, 1–12. [CrossRef] [PubMed]
14. Sun, Y.; Peng, M.; Zhou, Y.; Huang, Y.; Mao, S. Application of machine learning in wireless networks: Key techniques and open issues. *IEEE Commun. Surv. Tutor.* 2019, 21, 3072–3108. [CrossRef]
15. 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]
16. Elbir, A.M.; Papazafeiropoulos, A.; Kourtessis, P.; Chatzinotas, S. Deep channel learning for large intelligent surfaces aided mm-wave massive MIMO systems. *IEEE Wirel. Commun. Lett.* 2020, 9, 1447–1451. [CrossRef]
17. Shtaiwi, E.; Zhang, H.; Abdelhadi, A.; Han, Z. RIS-Assisted mmWave Channel Estimation Using Convolutional Neural Networks. In *Proceedings of the 2021 IEEE Wireless Communications and Networking Conference Workshops (WCNCW), Nanjing, China, 29 March 2021*, pp. 1–6. [CrossRef]
18. Coleri, S.; Ergen, M.; Puri, A.; Bahai, A. Channel estimation techniques based on pilot arrangement in OFDM systems. *IEEE Trans. Broadcast.* 2002, 48, 223–229. [CrossRef]
19. 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]
20. Taha, A.; Alrabeiah, M.; Alkhateeb, A. Enabling Large Intelligent Surfaces With Compressive Sensing and Deep Learning. *IEEE Access* 2021, 9, 44304–44321. [CrossRef]

21. 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]

22. Kiranyaz, S.; Avci, O.; Abdeljaber, O.; Ince, T.; Gabbouj, M.; Inman, D.J. 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* 2021, 151, 107398. [CrossRef]

23. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the 32nd International Conference on Machine Learning (ICML), Lille, France, 6–11 July 2015; Volume 37, pp. 448–456. [CrossRef]

24. 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]