International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Super-Resolution Channel Estimation based Deep Learning in Reconfigurable Intelligent Surface Systems

Wala'a Hussein[1 *], Kamil Audah[2], Nor K. Noordin[3], Mod Fadlee B. A. Rasid[4], Alyani Binti Ismail[5]

Submitted: 04/12/2023 **Revised**: 13/01/2024 **Accepted**: 25/01/2024

Abstract: The propagation environment may be configured using a reconfigurable intelligent surface (RIS). The channel estimate is a critical problem in implementing the RIS-aided communication system. A cascaded channel with large dimensions and complex statistics is used in a RIS-aided multi-user multiple-input multiple-output (MIMO) orthogonal frequency division multiplexing (OFDM) communication system. This research elucidates the application of deep learning (DL) for communication channel estimation. It employs a two-dimensional visualization to represent the time-frequency attributes of a rapidly fading communication channel. The objective is to decipher the undisclosed channel response by contrasting it with recognized values at designated "pilot points." We have introduced an extensive framework that integrates sophisticated image processing methods, including techniques like image super-resolution (SR) and image restoration (IR), to accomplish this goal. This method views the pilot data collectively as a low-quality image and calculates the channel using an SR system paired with a noise-reducing IR system.Moreover, a practical application of the proposed procedure is also detailed. According to the simulation results, the trained DL estimator outperforms the Least Square estimators in predicting the channel and identifying transmitted symbols, although the suggested SRIR estimator is more sophisticated. Furthermore, the DL estimator exhibits its efficacy with varying pilot densities and cycle prefix times. The findings show that this pipeline may be utilized effectively in channel estimation.

***Keywords**: Channel estimation, Deep Learning, Image Super-resolution, Image restoration, MIMO-OFDM, RIS.*

1. Introduction

The sixth-generation (6G) wireless networks would allow for high-speed data transfer rates, extremely low latency, many simultaneous connections, and fantastic mobility. Database applications in cellular networks have expanded greatly over the last five years [1]. These include multimedia, online entertainment, and HD video streaming. As a result, both the use of mobile devices and the speed with which information can be sent have exploded in recent years. Consequently, new research domains present an array of innovative solutions, including massive multiple input multiple outputs (mMIMO), reconfigurable intelligent surfaces (RIS), and millimetre Wave (mmWave) communications [2]. These advancements aim to address the anticipated surge in wireless data traffic and ensure enhanced reliability and security for the forthcoming generation. OFDM is a modulation technique widely adopted in communication systems to counteract frequency-selective fading in wireless networks. Frequently, the signal received through a communication channel is distorted by the channel's inherent properties. The channeling effect must be evaluated and adjusted at the receiver to retrieve the broadcast symbols. Generally, the receiver estimates the channel using certain symbols known as pilots, whose locations and time-frequency values are known to both the transmitter and the receiver. Three distinct structures may be proposed based on these pilot arrangements: block-type, comb-type, and lattice-type [1]. Pilots are broadcast on all available subcarriers at the beginning of each OFDM block. In contrast, in a comb-type configuration, pilots occur on many subcarriers associated with individual OFDM signals. Lattice-type designs have pilots spaced in a diamond pattern across the time and frequency dimensions. Conventional methods that utilize pilots for estimation, such as Least Square (LS) and Minimum Mean Square Error (MMSE), leverage pilot values within time-frequency matrices to ascertain undetermined channel responses. These methods have been refined across multiple scenarios [2]. Unlike the LS estimation that operates without channel statistical data, MMSE enhances accuracy by utilizing channel statistics and noise variance. To implement MMSE in practice, we discuss concrete strategies for simplifying the system by making do with an approximation of the channel statistics rather than the real item. Approximate Linear MMSE (ALMMSE) is a simplified variant of the MMSE established in [3] for fast-fading channels. Since the correlation and filtering matrices have been significantly reduced in size, the overall

Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia [1,2,3,4,5];
Wireless and Photonics Networks Research Centre of Excellence (WiPNET), Faculty of
Engineering, Universiti Putra Malaysia, Serdang, Selangor 43400, Malaysia [1,2,3]
Department of Computer Technology Engineering, Faculty of Engineering, Iraq University College, Basra, Iraq [1,2] walaahussein613@gmail.com
Department of Chemical Engineering and Petroleum Refining Basrah University for Oil and Gas, Basra, Iraq [1]
Corresponding author: Wala'a Hussein (walaahussein613@gmail.com), Nor K. Noordin (nknordin@upm.edu.my).

complexity of this method is much lower than that of the standard MMSE. Lately, the field of communication systems has been more interested in deep learning (DL). There have been introductions of DL-based methodologies to enhance the efficiency of various standard algorithms, such as modulation identification [4, 5], signal pinpointing [6, 7], channel equalization [8, 9], and channel assessment [8]. The communication system is perceived as an opaque entity [8], with a comprehensive DL framework being employed for both signal sending and receiving. This DL section intrinsically integrates encoding, decoding, channel evaluation, and other fundamental communication link processes. Moreover, applications that need the whole channel response will not benefit from this technique because of its capacity to directly assess the channel's time-frequency response. The approach outlined in [9] utilizes an image-based depiction of the channel matrix within a denoising network for approximating the channel. This study predominantly concentrates on the channel matrix in the transmitter/receiver antenna domains, avoiding an in-depth analysis of the time-frequency response for individual Tx/Rx links, particularly in multi-antenna configurations. The Reconfigurable Intelligent Surface (RIS) emerges as a promising tool for advancing intelligent radio landscapes in next-generation systems [24], owing to its capability to shape wireless signal propagation via its extensive array of passive elements. Leveraging this, RIS can amplify key performance indicators such as throughput, coverage, energy efficiency (EE), and spectral efficiency (SE) within wireless communication systems [23]. To leverage the potential improvements in Reconfigurable Intelligent Surface (RIS)-supported wireless communication systems, having precise Channel State Information (CSI) is crucial, especially for implementing precoding. However, unlike traditional systems, RIS introduces a composite BS-RIS-user channel that doesn't follow the standard Gaussian distribution. In such cases, determining the optimal Minimal Mean Square Error (MMSE) estimator becomes infeasible. Additionally, while the Least Square (LS) and Linear MMSE (LMMSE) estimators are employed, a performance gap remains compared to the ideal MMSE estimator. The accuracy of their estimates falls short of meeting the requirements for establishing an effective RIS-aided communication system. Moreover, due to the large number of reflecting elements in an RIS and the multi-dimensional nature of the cascaded channel, CSI estimation using LS and LMMSE estimators demands substantial computational resources. Given the previously mentioned challenges, there has been a focused shift towards adopting deep learning (DL)-based methods for channel estimation in communication systems integrated with Reconfigurable Intelligent Surfaces (RIS), as detailed in references [25-27]. Amidst a multi-user environment enhanced by RIS, the contributions by Liu et al. stand out. Specifically, [25] used a deep convolutional residual neural network to improve the channel matrix derived via the LS technique. Kundu and colleagues [26] introduced two distinct convolutional neural Network (CNN) methods to enhance channel fidelity in a RIS-supported single-user MISO setup: namely, the denoising CNN (DnCNN) and the agile denoising network (FFDNet). Furthermore, Liu and team [27] put forth a complex-valued denoising convolutional neural Network to bolster the effectiveness of compressive sensing-driven channel estimation in a RIS-integrated mmWave extensive MIMO environment. In these research efforts, the channels enhanced by RIS estimation form essential input for the neural network processes. Estimating RIS-enhanced channels poses computational challenges, chiefly due to the dimensionally expansive nature of the cascaded channel.

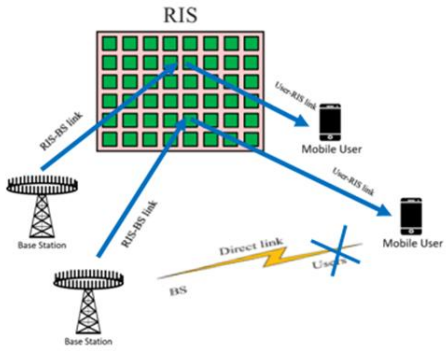


Fig. 1.RIS-aided downlink wireless communication.

The key contributions of this article can be outlined as:

- Represent the channel's time-frequency response in a two-dimensional image utilizing the TDL-A_45ns model. We employed the Vienna 5G Link Level simulator for channel modeling and pilot transmission, which the University of Vienna developed and operated on MATLAB 2022a. For the instantiation and elaboration of our proposed framework, we utilized both Keras and TensorFlow, leveraging a GPU backend for computational efficiency.
- In contrast to existing research, we recast the channel estimation challenge in a RIS-supported multi-user MIMO-OFDM communication framework as an image super-resolution (SR) issue. Stemming from this perspective, we introduce a deep CNN-driven methodology, SRIR-ChNet, designed to reconstruct the channel matrix by drawing estimations from channels at designated pilot positions. Owing to the advanced interpolation and noise reduction capabilities inherent in the refined SRCNN and DnCNN, the SRIR-ChNet is poised to enhance the precision of channel estimation significantly.
- We put forth two algorithms by interpreting the channel response at pilot points as a Low-Resolution (LR) image and considering the anticipated channel response as the proposed High-Resolution (HR) image. Utilizing deep learning, the SRCNN and DnCNN are designed to

enhance the resolution of the Time/frequency response channel matrix image and mitigate its noise. The proposed algorithm can achieve high accuracy in terms of channel estimation compared with conventional methods by minimizing mean square error and bit error rate.

- Furthermore, the impact of the quantity of pilots on the performance of SRIR-ChNet in the context of a MIMO-OFDM network is shown.

The subsequent sections of the paper are structured in the following manner: Channel estimation using traditional techniques is briefly covered in Section II. The suggested DL-based channel estimator and its motivations and design are presented in Section III. Our proposed scheme's network architecture V, simulation results, and analysis are described in part V, and the article concludes to a conclusion in section VI.

2. Background

A. CHANNEL ESTIMATION

Channel estimation (CE) is one of the primary methods used in Orthogonal Frequency Division Multiplexing Modulation in cellular wireless communication systems (OFDM). The most often used approaches are decision-directed channel estimation (DCE), pilot-assisted channel estimation (PACE), and blind channel estimation. PACE is the most frequently utilized and has a more consistent performance. Deep learning (DL) approaches have piqued the attention of scholars during the last three years. In an OFDM system, the input-output connection for the k_{th} time slot and i_{th} subcarrier is expressed as [17]:

$$Y_{i,k} = H_{i,k} X_{i,k} + Z_{i,k}. \quad (1)$$

Consider an OFDM subframe sized $N_S \times N_D$. The time slot index, denoted by k, ranges from $[0, N_D - 1]$, and the subcarrier index, represented by i, stretches from $[0, N_S - 1]$. As per equation (1), $Y_{i,k}$, $X_{i,k}$, and $Z_{i,k}$ represent the signal received, the transmitted OFDM symbol, and the white Gaussian noise, respectively. The element $H_{i,k}$ is a specific (i, k) entry of matrix H, which belongs to the set $C\ N_S \times N_D$. H captures the channel's time-frequency response for every subcarrier and time slot. In the realm of channel estimation, particularly in channels subject to fading, the representation in the time domain is specified as H = {h[1], h[2], ..., h[N_D]}. In this representation, each h[k] denotes the channel's frequency response during the kth time slot. The LS method is employed to determine the channel at the pilot points. When we view the LS-estimated channel as a diagonal matrix, denoted $\boldsymbol{H}_P^{LS} \in C^{NP \times NP}$, it's inferred from [17] that the form of $\boldsymbol{H}_P^{LS}$ is:

$$\widehat{\boldsymbol{H}}_P^{LS} = arg_{\boldsymbol{H_P}} \min ||y_P \quad - \boldsymbol{H}_P x_P||_2^2, \quad (2)$$

where $||.||2$ is the $\mathscr{L}2$ distance and $\widehat{\boldsymbol{H}}_P^{LS} \in C^{NP \times NP}$ is the estimated diagonal matrix. x_p contains the known pilot values, and y_p is the corresponding observations. Solving optimization (2) yields in $\widehat{\boldsymbol{h}}_P^{LS}$ = diag($\widehat{\boldsymbol{H}}_P^{LS}$) = y_p/x_p. We need to use a bidimensional interpolation technique to identify the channel value at locations other than pilot points. When compared to LS, the MMSE estimator is the superior choice. Multiplying the LS estimates at the pilot-symbol positions by a filtering matrix yields this AMMSE $\in C^{NL \times NP}$ as mentioned in [14]:

$$\widehat{\boldsymbol{h}}_d^{MMSE} = \mathbf{A}_{\text{MMSE}} \widehat{\boldsymbol{h}}_P^{LS}, \quad (3)$$

Where $\widehat{\boldsymbol{h}}_d^{MMSE} = \in C^{NL \times 1}$ ($N_L = N_S \times N_D$) is the linearized MMSE approximation of the channel response H for subframe d. To determine the filtering matrix, one considers the mean square error (MSE),

$$€ = E\{||\mathbf{h}d - \mathbf{A}_{\text{MMSE}} \widehat{\boldsymbol{h}}_P^{LS}||_2^2\}, \quad (4)$$

has to be minimized. Minimizing (4) leads to

$$\mathbf{A}_{\text{MMSE}} = R R_{h_{dh_p}} (\boldsymbol{R}_{h_P h_P} + \sigma_n^2 (XX^H)^{-1})^{-1}, \quad (5)$$

In this context, the matrix $\boldsymbol{R}_{h_P h_P} = E\{h_d h_p^H\}$ denotes the channel correlation matrix correlating the desired subframe with pilot symbols. Similarly, the matrix $\boldsymbol{R}_{h_P h_P} = E\{hd h_p^H\}$ describes the channel correlation matrix exclusive to the pilot symbols. For MMSE to function optimally, it's imperative that the complete correlation matrix of the channel symbolized as R, is comprehensively known.

B. SUPER-RESOLUTION AND IMAGE RESTORATION

Various methods have been devised to enhance pixelated, low-resolution images into sharper, high-resolution counterparts. The umbrella term for these methods is "image super-resolution" (SR), designed specifically for image enhancement. Deep learning algorithms, particularly those utilizing intricate and layered convolutional networks, demonstrate proficiency in converting low-resolution images into high-definition versions. One notable example is the Super-resolution Convolutional Neural Network (SRCNN) [10]. Recently introduced, it serves as evidence of this advancement. Alongside super-resolution (SR) techniques, strategies for Image Restoration (IR) have been formulated to reduce or eliminate image noise. Numerous IR models have been extensively documented in academic literature. For example, the study in [11] unveils a methodology based on a feed-forward denoising convolutional neural Network (DnCNN). This approach incorporates residual learning with batch normalization to expedite the training process.

3. Srir-Chnet

C. RELATED WORKS AND MOTIVATION

A prominent research focus in image processing has been transforming low-resolution and noisy images into more precise, high-resolution versions. Image Super-Resolution (SR) encompasses a collection of methods dedicated to

refining the resolution of digital images. Deep CNNs have proven effective in extracting HR images from LR photos [14, 17]. Recently, it has been proposed that a super-resolution convolutional neural network (SRCNN) be used to map LR pictures completely [18, 19]. In addition to super-resolution techniques, various image restoration (IR) models, such as [20], [21], and [22], have been documented in the literature, aimed at mitigating the impact of noise on images. References [16, 23] have further proposed methods to accelerate the training of feed-forward denoising convolutional neural networks (DnCNNs) using batch normalization and residual learning.

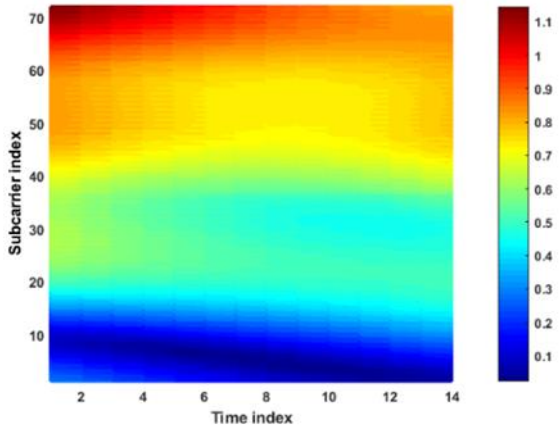


Fig. 2. An illustration of a normalized two-dimensional image showcasing real and imaginary components for a representative channel time-frequency grid.

This research envisions the channel response's time-frequency grid as a two-dimensional image and employs a sequential combination of the SR and IR networks for channel estimation. Figure 2 provides a normalized depiction of the time-frequency grid for a multipath Rayleigh fading channel, showcasing its real and imaginary components. The objective involves using broadcast pilots to infer the channel's complete time-frequency spectrum. The pattern chosen for pilot transmission echoes the configuration prevalent in the LTE standard.

B. ARCHITECTURE AND TRAINING

The presented pilot value of the channel, denoted as h, functions as a low-resolution and noise-affected channel depiction. This image is fed into the SR network, which interprets this noisy, low-resolution representation and vectorizes the channel estimate. Given that a picture's vector representation might encompass real and imaginary components, the SR network operates in two distinct phases: initially focusing on the natural aspect, followed by addressing the imaginary portion. The principal aim of the SR network is to ascertain the channel response, represented as h ~ n. This SR network is subsequently paired with an IR network to mitigate the influence of noise. In this context, the channel matrix is visualized as an image, and channel estimation is achieved through a sequenced SR and IR network approach, viewing the channel response's time-frequency grid as a two-dimensional image. A normalized 2D representation of a multipath Rayleigh fading channel's time-frequency grid is showcased in Figure 2. The goal is to calculate an approximation of the channel's 24/7 frequency from the sent pilots. An arrangement that was identical to the LTE standard was adopted for the pilot broadcast.

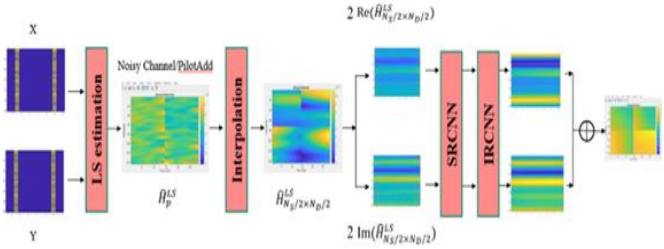


Fig. 3.A general pipeline of our Network

In the SRIR_ChNet framework, the SR network starts by using an approximation method to identify near-accurate values of an image, eventually generating a low-resolution depiction that aligns closely with the high-resolution original. The central objective is to create a mapping, denoted as F, incorporating a feature extractor, nonlinear mapping, and reconstruction, as detailed in [14]. In the initial phase, patches extracted from a low-resolution image are transformed into high-dimensional vectors. The length of which corresponds to the number of feature mappings used. This is followed by a nonlinear transition that maps one high-dimensional vector onto another, creating a new set of feature maps. The final step consolidates all these high-resolution patch representations to produce an image that is perceived to resemble the original closely. Figure 3 illustrates the integrated workflow of the SRIR network, harmonizing the stages above within a Convolutional Neural Network (CNN) structure.

1) Image network

In this work, attention is directed towards four communication links bridging Tx and Rx antennas, signifying the presence of multiple-input, multiple-output (MIMO) communication pathways. For such connections, the complex-valued channel time-frequency response matrix, H (of dimensions $N_S \times N_D$), between a transmitting and receiving entity can be delineated as a pair of two-dimensional images, encapsulating the fundamental values and the imaginary components. An exemplar of a normalized two-dimensional image, encompassing both real and fictional elements for a prototypical channel time-frequency grid with N_D equating to 14-time slots and N_S amounting to 72 subcarriers (based on 5G standards), is illustrated in Fig.2.

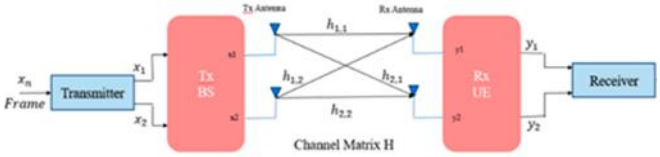


Fig. 4. Show 2x2 MIMO communication system.

OFDM-MIMO frame will be divided into four subframes (4 images), h_{11}, h_{12}, h_{21}, *and* h_{22}, and each subframe consists of 14-time slots with 72 subcarriers (1 photo).

$$x_1 = 72 \text{ x } 14, x_2 = 72 \text{ x } 14$$

$$y_1 = x_1 \text{ x } h_{11} + x_2 \text{ x } h_{12}$$

$$y_2 = x_1 \text{ x } h_{21} + x_2 \text{ x } h_{22}$$

$$H = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}$$

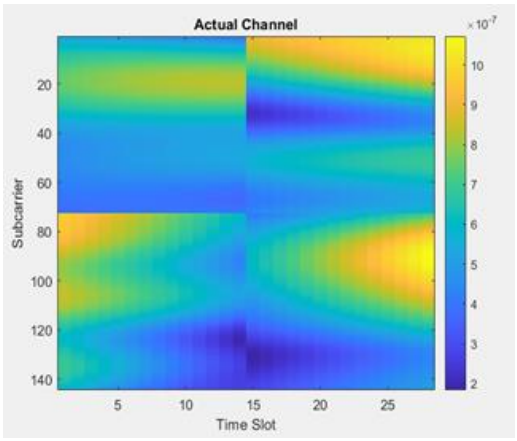


Fig. 5. One frame at SNR 10 has four 2D images due to 2 x 2 MIMO.

2) Network Structure

Figure 3 provides a schematic representation of the suggested DL-based channel estimation approach, SRIR-ChNet. The objective is to fully capture the channel's time-frequency spectrum through the broadcasted pilots. A Lattice-patterned framework, reminiscent of the LTE standard, was employed for the pilot transmissions. The representation of the channel, noted for its low fidelity and potential interference, is equivalent to the anticipated channel values at the pilot points, denoted as $\widehat{\ }$ HPLC. To achieve a comprehensive channel image, a two-phase training approach is introduced:

• In the initial phase, an SR network is formulated. This Network leverages $\widehat{h}_p^{LS}$ as the vectorized low-resolution input—processing the actual component first, followed by the imaginary segment—to deduce the unidentified channel response H values.

• In the subsequent phase, a noise-reducing IR network is combined with the SR network to mitigate noise interferences.

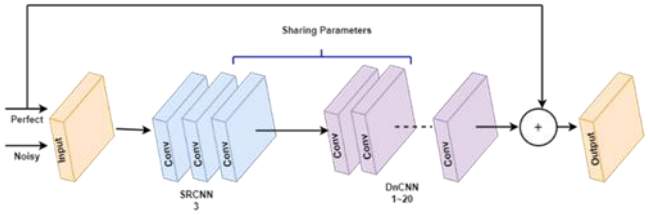


Fig. 6.SRIR-ChNet network structure.

We utilized SRCNN [10] for the SR and DnCNN [11] for the IR functions. Due to space limitations, we can't provide their graphical representations. Nonetheless, SRCNN uses a tri-layer convolutional network for enhancement after applying an interpolation method to roughly gauge high-resolution image (or channel) values. Post ReLu activation, the inaugural convolutional layer uses 64 filters, each measuring 9 x 9, while the subsequent layer employs 32 filters of 1 x 1 dimensions. The terminal layer reassembles the image through a singular 5 x 5 filter. As for DnCNN, it's a network based on residual learning and comprises 20 convolutional layers (further details can be found in [11]). The premier layer employs 64 filters, each with dimensions of 3 × 3 × 1, and is followed by a ReLU. The subsequent 18 convolutional layers utilize 64 filters of 3 x 3 x 64 dimensions, succeeded by batch-normalization and ReLU activations. The concluding layer uses a singular 3 x 3 x 64 filter to reconstruct the end output.

TABLE I. SETTINGS OF SRIR-CHNET

Input	SRCNN	IRCNN	Output
	Layers ***Filter Size***	***Layers*** ***Filter Size***	
Estimated channels at pilot positions.	1 64 × (9 × 9 × 2) 2 32 × (1 × 1 × 64) 3 2 × (5 × 5 × 32)	1 64 × (3 × 3 × 2) 2~19 64 × (3 × 3 × 64) 20 2 × (3 × 3 × 64)	Estimated whole channel.

4. SRIR-Chnet for MIMO-OFDM Network

As illustrated in Fig. 1, we analyze a downlink MIMO-OFDM system with a re-configurable intelligent surface (RIS) and multiple antennas (M) at the base station (BS) broadcasting to a large number of multiple-antenna users. In a scenario where the Reconfigurable Intelligent Surface (RIS) serves two Base Stations (BS) and K users, the BS is equipped with an NBS-antenna array, and the RIS has NM reflecting elements catering to both the BS and K multiple-antenna users. The BS antenna arrays and RIS reflecting components are arranged in a Homogeneous Linear Array (ULA). The BS controls a controller responsible for reconfiguring the propagation environment as illustrated in Fig. 4, h_1, $h_{2,k}$, and d_k, where $k \in \{1, 2, \ldots, K\}$, represent

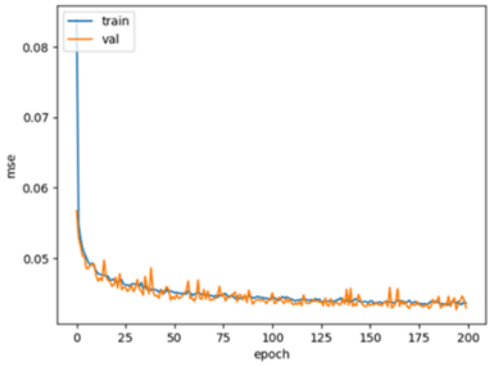
the channels of RIS-BS, U_k-RIS, and U_k-BS, respectively. Due to numerous scatters in the channel between the BS and users, the Line of Sight (LoS) path may not be present [37]. Therefore, a Rayleigh fading channel model is adopted to represent dk ∈ C NM×1. The channels h_1, h2, and $_k$ are modelled as Rician fading channels. Additionally, the system employs a Frequency-Division Duplex (FDD) mode to utilize channel reciprocity (CR). The Base Station (BS) can predict downlink channels through FDD based on the Channel Reciprocity (CR) principle. This involves setting channels to receive downlink pilots. Leveraging the channel estimation results, the BS can then implement Zero-Forcing (ZF) Beamforming (BF). This study focuses explicitly on Zero-Forcing (ZF)-BF, aiming to mitigate intra-cluster and inter-cluster interference [33]. BF is a favoured technique in MIMO networks, especially when the BS is equipped with multiple antennas.

5. Simulation Results and Analysis

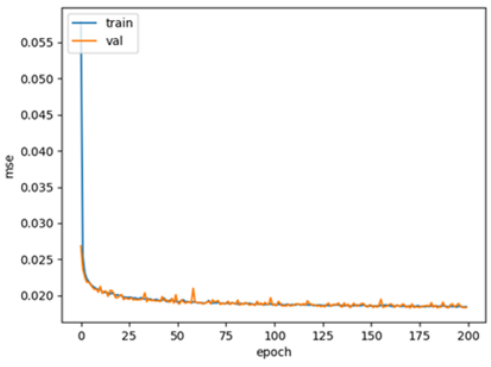
In our research, we employ link-level simulation to evaluate the efficacy of the SRIR-ChNet network's channel estimation in a MIMO-OFDM environment. In the 5G framework, a single time slot within the time-frequency dimension consists of 14 OFDM symbols. On the frequency side, there's an aggregate of 72 subcarriers. These are divided into 6 Resource Blocks (RBs), with each RB containing 12 subcarriers. A channel simulator from the University of Vienna [21] produces the TDL-A 45ns channel model. This configuration is defined by parameters such as a frequency of 2.1 GHz, a bandwidth of 1.6 MHz, a delay spread of 1730 ns, and a 50 km/h velocity for the user equipment (UE). We test the effectiveness of our suggested method over a spectrum of signal-to-noise ratio (SNR) scenarios. Different noisy channel data sets, ranging between 0 and 30 dB, are generated using the TDL-A 45ns channel model to train and evaluate the SRIR-ChNet network. The training and testing datasets consist of 50,000 and 10,000 channel realizations. This research adopts the mean square error (MSE) as its primary assessment metric to juxtapose the forecasted channel value with its counterpart.

TABLE II. THE PARAMETER SETUP FOR THE CONSIDERED MIMO-OFDM SYSTEM.

Parameters	Values
MIMO, RIS, BS, UE	2 x 2 , 1,2,2
Distance of RIS-BS	100m
Simulation	Downlink
The Nearest Distance between users and BS	102m
Width of Reflecting Elements	0.012m
Length of Reflecting Elements	0.012m
Type of modulation	OFDM 64 QAM
Maximum Doppler frequency	36 Hz, 200 Hz
Noise model	Gaussian Noise
Sample frequency	3.84 MHz
No. of Sub-carriers	72
Time Slot	14
Fading	Rayleigh
Frame Structure	FDD
Channel Power Delay Profile	TDL-A_45ns
Channel Estimation Method	Approximate-Perfect, PilotAided LS
Pilot Pattern Downlink	Diamond

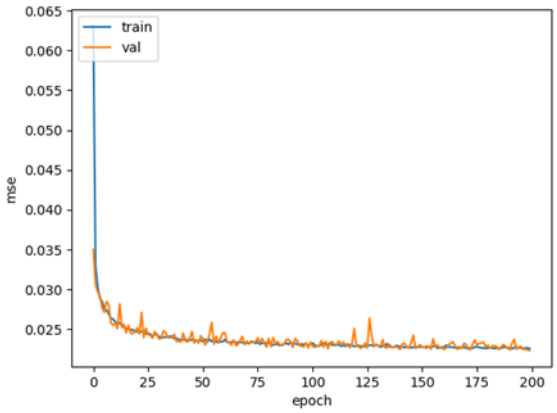


(a)

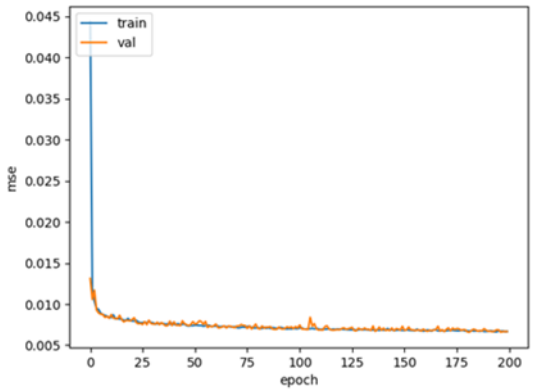


(b)

Fig. 7.MSE channel estimation for training and validation (a)SNR=10 & pilot=8(b) SNR=20 & pilot=8

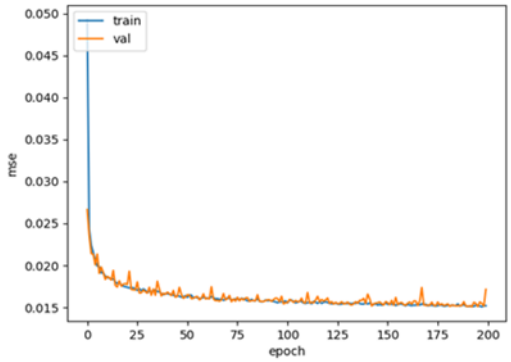


(a)

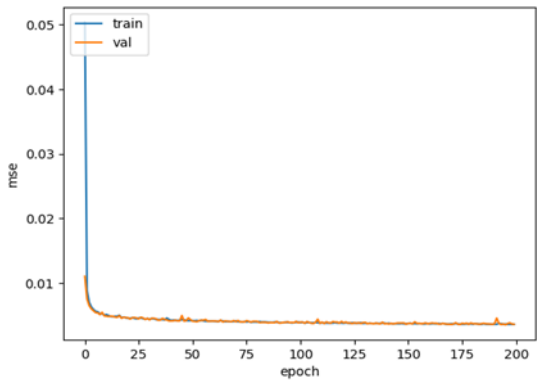


(b)

Fig. 8.MSE channel estimation for training and validation (a)SNR=10 & pilot=16(b) SNR=20 & pilot=16

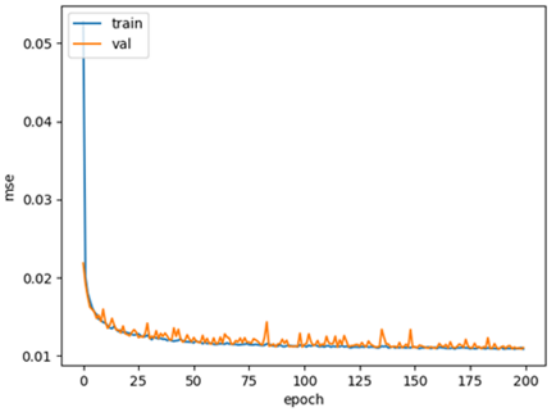


(a)

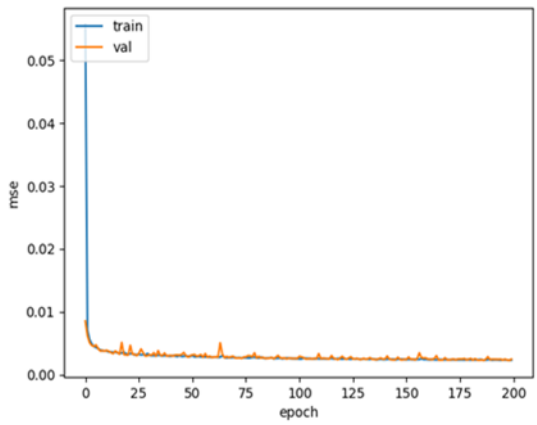


(b)

Fig. 9.MSE channel estimation for training and validation (a)SNR=10 & pilot=24(b) SNR=20 & pilot=24

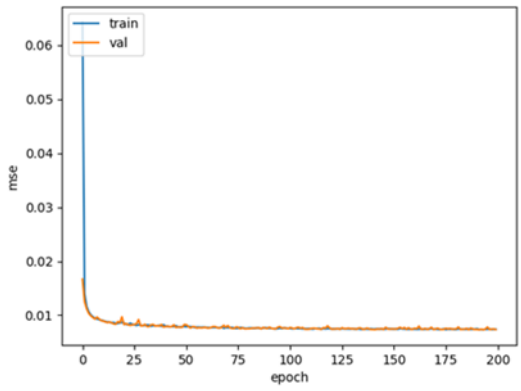


(a)

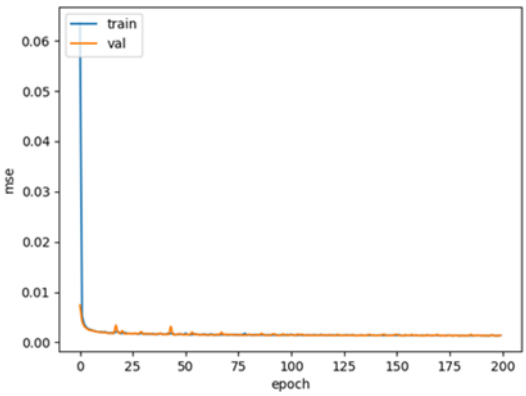


(b)

Fig. 10. MSE channel estimation for training and validation (a)SNR=10 & pilot=36(b) SNR=20 & pilot=36



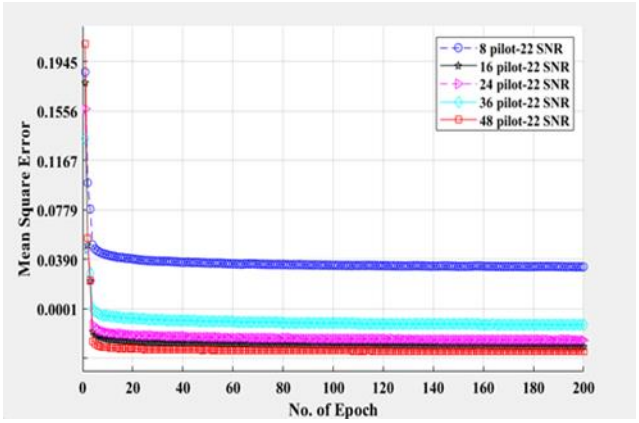
(a)



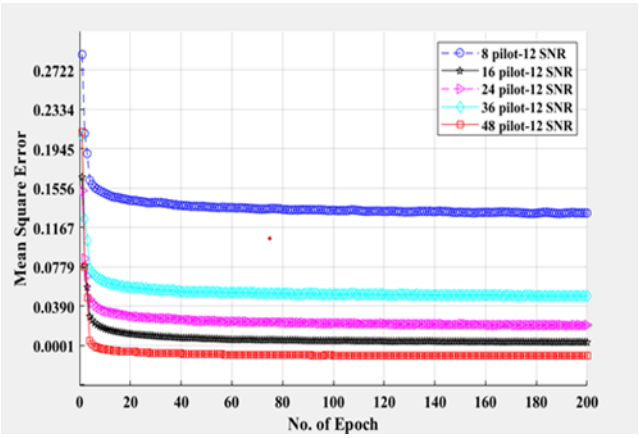
(b)

Fig. 11. MSE channel estimation for training and validation (a)SNR=10 & pilot=48(b) SNR=20 & pilot=48

Simulation outcomes reveal that the SRIR-ChNet methodology exhibits superior estimation accuracy compared to the traditional technique using uniformly spaced pilots. As illustrated in Figure 12, the enhancement in channel estimation precision due to pilot optimization becomes more evident with fewer pilots. Beyond merely elevating channel estimation precision, pilot design plays a pivotal role in curtailing pilot overhead. In this analysis, we assessed channel estimation precision by altering the pilots' count while maintaining a consistent signal-to-noise ratio. The accuracy of each algorithm invariably advances as pilot numbers swell. Given that the optimal MMSE algorithm exclusively predicts channel data of the specific pilot locations it occupies, its estimation precision remains relatively stable, regardless of pilot count variations. With eight pilots in place, the SRIR-ChNet methodology advocated in this study surpasses the traditional LS algorithm regarding estimation accuracy. At a pilot count of forty-eight, the SRIR-ChNet approach necessitates fewer pilots compared to the MMSE method to achieve comparable estimation precision. As a result, the introduced SRIR-ChNet methodology possesses significant promise in diminishing pilot overhead during channel estimation. Beyond the influence of pilot positioning on accuracy (specifically, the decision to adopt an optimal pilot arrangement), the total count of unique pilots remains paramount. Figure 12 offers a comparative analysis of the accuracy achieved by SRIR-ChNet in channel estimation for pilot counts of 8, 16, 24, 36, and 48. There's a proportional rise in estimation accuracy with increased pilot numbers. For SNR values greater than 18 dB, the estimation precision for a time-frequency grid comprising 1008 Resource Elements (RE) remains consistent between 16 and 48 pilots, specifically at an SNR of 22. Remarkably, accuracy levels can converge between magnitudes of 10-3 and 10-4, even with just eight pilots. This underscores that, at heightened signal-to-noise ratios, the SRIR-ChNet approach exhibits robust adaptability to variations in pilot numbers. Moreover, it holds the potential to elevate the system's spectral efficiency while safeguarding the integrity of the channel estimation accuracy.



(a)



(b)

Fig. 12. Mean square error for channel estimation in terms of pilot number(a) at SNR 12 (b) at SNR 22

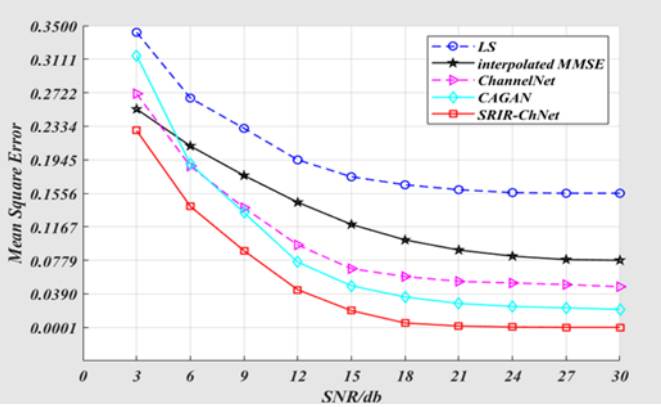


Fig. 13. Performance comparison of different algorithms when the number of pilots is 8.

In Figure 13, we evaluate the efficiency of various channel estimation methods. Our SRIR-ChNet is juxtaposed against four techniques mentioned in [22]: namely, least squares (LS), interpolated minimum mean square error (interpolated MMSE), conditional generative adversarial network (CAGAN), and ChannelNet. A consistent experimental framework is applied across all tests. As depicted in Figure 13, the SRIR-ChNet introduced in this study outperforms the conventional LS and MMSE techniques, underscoring the advantages of employing deep learning approaches. When the SNR surpasses 15 dB, the estimation prowess of all methods stabilizes. However, with estimation precision for ChannelNet ranging from 10-1 to 10-2, it needs to catch up to the superior performance of SRIR-ChNet. When the number of pilots is increased to 8, under the same SNR condition, SRIR-ChNet outperforms the other four methods; estimate accuracy is improved to between 10-3 and 10-4, and SRIR-ChNet and CAGAN are considerably closer under high signal-to-noise ratio.

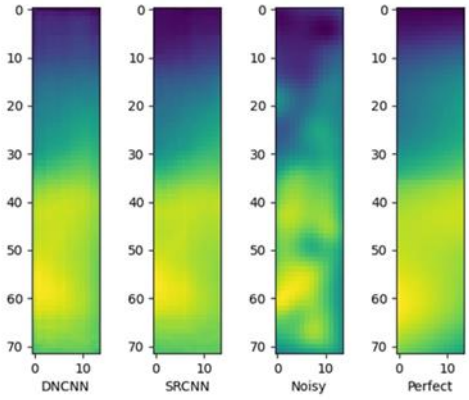


Fig. 14. The output of SRIR-ChNet 2D image time-frequency response.

6. Conclusion

This study introduces a practical transmission protocol tailored for channel estimation and reflection optimization in RIS-augmented MIMO-OFDM systems. We showcase a deep learning model grounded in image representation for next-gen wireless communication channel estimation. This model employs a 2D visual depiction of a Rayleigh fading communication channel's time-frequency behaviour within a reconfigurable intelligent surface (RIS) integrated multi-user MIMO-OFDM setup. Our approach aims to discern the undisclosed channel response values by utilizing a comprehensive framework encompassing advanced image processing modalities, notably image SR and noise-reduction methodologies. This work advocates a deep learning-centric strategy to achieve precise channel predictions. We redefine channel estimation as an image super-resolution challenge and introduce the SRIR-ChNet, a novel image super-resolution network crafted to reconstruct the channel matrix from initial channel assessments. Leveraging the capabilities of the advanced super-resolution convolutional neural Network (SRCNN) and the denoising convolutional neural Network (DnCNN), the SRIR-ChNet enhances channel estimation capabilities regarding feature extraction and noise mitigation. We evaluated our proposed technique against various benchmarks, such as signal-to-noise ratios (SNRs) and pilot count. Simulation results, gauged using normalized mean square error metrics, confirm that SRIR-ChNet outpaces traditional channel predictors and DL-infused methodologies by a margin of over 10 dB. Our findings reveal that for satisfactory channel prediction performance under an SNR of 5dB, a minimum of 8 pilots is requisite. Moreover, SRIR-ChNet promises reduced input data volume and computational overhead tied to channel prediction in intricate scenarios. Future exploration could pivot towards refining the performance of the mentioned mMIMO-RIS structure by adopting innovative pilot pattern designs.

ACKNOWLEDGEMENTS

The authors are grateful to The University Putra Malaysia Faculty of Computer and Communication Engineering, Iraq University College.

References

[1] L. Yu, Z. Liu, M. Wen, D. Cai, S. Dang, Y. Wang, and P. Xiao, “Sparse Code Multiple Access for 6G Wireless Communication Networks: Recent Advances and Future Directions,” IEEE Communications Standards Magazine, 2021.

[2] B. Zhang, Y. Wang, W. Wang, and Y. Tian, “On the Downlink Throughput Capacity of Hybrid Wireless Networks With MIMO,” IEEE Access, vol. 5, pp. 26 086–26 091, 2017.

[3] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on the pilot arrangement in OFDM systems," IEEE Transactions on Broadcasting, vol. 48, pp. 223–229, Sep 2002.

[4] Y. Li, L. J. Cimini, and N. R. Sollenberger, "Robust channel estimation for OFDM systems with rapid

dispersive fading channels," IEEE Transactions on Communications, vol. 46, pp. 902–915, Jul 1998.

[5] M. imko, C. Mehlfhrer, M. Wrulich, and M. Rupp, “Doubly dispersive channel estimation with scalable complexity,” in 2010 International ITG Workshop on Smart Antennas (WSA), pp. 251–256, Feb 2010.

[6] T. OShea and J. Hoydis, “An introduction to deep learning for the physical layer,” IEEE Transactions on Cognitive Communications and Networking, vol. 3, pp. 563–575, Dec 2017.

[7] N. Samuel, T. Diskin, and A. Wiesel, “Deep mimo detection,” in 2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), pp. 1–5, July 2017.

[8] D. Erdogmus, D. Rende, J. C. Principe, and T. F. Wong, "Nonlinear channel equalization using multilayer perceptrons with information theoretic criterion," in Neural Networks for Signal Processing XI: Proceedings of the 2001 IEEE Signal Processing Society Workshop (IEEE Cat. No.01TH8584), pp. 443–451, 2001.

[9] C. Wen, W. Shih, and S. Jin, "Deep learning for massive mimo CSI feedback," IEEE Wireless Communications Letters, pp. 1–1, 2018.

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

[11] H. He, C. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mm-wave massive mimo systems," IEEE Wireless Communications Letters, vol. 7, pp. 852–855, Oct 2018.

[12] C. Dong, C. C. Loy, K. He, and X. Tang, “Image super-resolution using deep convolutional networks,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, pp. 295–307, Feb 2016.

[13] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, “Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising,” IEEE Transactions on Image Processing, vol. 26, pp. 3142–3155, July 2017.

[14] S. Omar, A. Ancora, and D. T. M. Slock, "Performance analysis of general pilot-aided linear channel estimation in LTE OFDMA systems with application to simplified MMSE schemes," in 2008 IEEE 19th International Symposium on Personal, Indoor, and Mobile Radio Communications, pp. 1–6, Sept 2008.

[15] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-resolution using Deep Convolutional Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295–307, 2015.

[16] W. Gao, M. Yang, W. Zhang, Y. Zhao, Y. Zhou, and K. Zhang, "FFDNet-Based Channel Estimation for Beamspace mmWave Massive MIMO Systems," in 2021 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data and Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom). IEEE, 2021, pp. 1167–1170.

[17] M. Soltani, V. Pourahmadi, A. Mirzaei, and H. Sheikhzadeh, “Deep Learning-Based Channel Estimation,” IEEE Communications Letters, vol. 23, no. 4, pp. 652–655, 2019.

[18] A. Naseer, T. Yasir, A. Azhar, T. Shakeel, and K. Zafar, "Computer-aided Brain Tumor Diagnosis: Performance Evaluation of Deep Learner CNN Using Augmented Brain MRI," International Journal of Biomedical Imaging, vol. 2021, 2021.

[19] C. Ledig, L. Theis, F. Husz´ar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. Tejani, J. Totz, Z. Wang, et al., "Photo-realistic Single Image Super-resolution using a Generative Adversarial Network," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4681–4690.

[20] W. Yang, X. Zhang, Y. Tian, W. Wang, J.-H. Xue, and Q. Liao, “Deep Learning for Single Image Super-resolution: A Brief Review,” IEEE Transactions on Multimedia, vol. 21, no. 12, pp. 3106–3121, 2019.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[22] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142–3155, 2017.

[23] Y. Zhao, Y. Li, X. Dong, and B. Yang, “Low-Frequency Noise Suppression Method Based on Improved DnCNN in Desert Seismic Data,” IEEE Geoscience and Remote Sensing Letters, vol. 16, no. 5, pp. 811–815, 2019.

[24] R. Lan, H. Zou, C. Pang, Y. Zhong, Z. Liu, and X. Luo, “Image Denoising via Deep Residual Convolutional Neural Networks,” Signal, Image and Video Processing, vol. 15, no. 1, pp. 1–8, 2021.

[25] M. Di Renzo, A. Zappone, M. Debbah, M. S. Alouini, C. Yuen, J. de Rosny, and S. Tretyakov, “Smart radio environments empowered by reconfigurable intelligent

surfaces: How it works, state of research, and the road ahead," IEEE J. Sel. Areas Commun., vol. 38, no. 11, pp. 2450–2525, Nov. 2020.

[26] C. Huang, A. Zappone, G. C. Alexandropoulos, M. Deb- bah, and C. Yuen, "Reconfigurable intelligent surfaces for energy efficiency in wireless communication," IEEE Trans. Wireless Commun., vol. 18, no. 8, pp. 4157–4170, Aug. 2019.

[27] C. Liu, X. Liu, D. W. K. Ng, and J. Yuan, "Deep Residual Learning for Channel Estimation in Intelligent Reflecting Surface-Assisted Multi-User Communications," IEEE Trans. Wireless Commun., 2021, doi=10.1109/TWC.2021.3100148.

[28] N. K. Kundu and M. R. McKay, "Channel Estimation for Reconfigurable Intelligent Surface Aided MISO Communications: From LMMSE to Deep Learning Solutions," IEEE Open J. Commun. Soc., vol. 2, pp. 471–487, 2021.

[29] S. Liu, Z. Gao, J. Zhang, M. D. Renzo, and M.- S. Alouini, "Deep Denoising Neural Network Assisted Compressive Channel Estimation for mmWave Intelligent Reflecting Surfaces," IEEE Trans. Veh. Technol., vol. 69, no. 8, pp. 9223–9228, Aug 2020.