Lightweight Convolutional Neural Networks for CSI Feedback in Massive MIMO

Zheng Cao, Wan-Ting Shih, Jiajia Guo, Chao-Kai Wen, *Member, IEEE*, Shi Jin, *Senior Member, IEEE*

Abstract—In frequency division duplex mode of massive multiple-input multiple-output systems, the downlink channel state information (CSI) must be sent to the base station (BS) through a feedback link. However, transmitting CSI to the BS is costly due to the bandwidth limitation of the feedback link. Deep learning (DL) has recently achieved remarkable success in CSI feedback. Realizing high-performance and low-complexity CSI feedback is a challenge in DL based communication. We develop a DL based CSI feedback network in this study to complete the feedback of CSI effectively. However, this network cannot be effectively applied to the mobile terminal because of the excessive numbers of parameters. Therefore, we further propose a new lightweight CSI feedback network based on the developed network. Simulation results show that the proposed CSI network exhibits better reconstruction performance than that of other CsiNet-related works. Moreover, the lightweight network maintains a few parameters and parameter complexity while ensuring satisfactory reconstruction performance. These findings suggest the feasibility and potential of the proposed techniques.

Index Terms—Massive MIMO, FDD, CSI feedback, deep learning, lightweight neural network.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems are widely regarded as the main technology of 5G wireless communication systems [1]. These systems refer to a communication system that uses hundreds of antennas at the same time-frequency resource to serve tens of user equipment (UE) simultaneously. In frequency division duplex (FDD) mode, a massive MIMO system must feedback the downlink channel state information (CSI) to the base station (BS) to realize its potential gain. However, transmitting CSI to the BS is costly due to the bandwidth limitation.

In recent years, deep learning (DL) technology has been widely applied in the field of wireless communication [2], [3]. The authors in [4] used DL technology to build a new type of CSI sensing and recovery neural network called CsiNet. This network can learn how to use the channel architecture effectively for the conversion from CSI to codeword and vice versa. The reconstruction performance of CsiNet is superior to that of traditional compressed sensing methods. Subsequent related studies expanded the original scope of the network. [5] considered the time correlation to improve the reconstruction performance of the network. [6] utilized the correlation between the up-link and downlink to reduce the

Z. Cao, J. Guo and S. Jin are with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, P. R. China (e-mail: 707647403@qq.com, {jiajiaguo, jinshi}@seu.edu.cn).

W.-T. Shih and C.-K. Wen are with the Institute of Communications Engineering, National Sun Yat-sen University, Kaohsiung 80424, Taiwan (e-mail: sydney2317076@gmail.com, chaokai.wen@mail.nsysu.edu.tw).

CSI feedback payload remarkably. [7] used a deep recurrent neural network to learn temporal correlation and adopted deep separable convolution to shrink the model, thereby improving the reconstruction performance. [8] minimized the effect of feedback transmission errors and delays. [9] proposed a novel quantization CSI feedback network that considers non-uniform μ-law quantization to obtain uniformly distributed quantization symbols. [10] introduced a quantization-based entropy coding with an architecture comprising convolutional layers, followed by quantization and entropy coding blocks. [11] proposed a DL-based multi-user collaborative feedback architecture, which is a distributed feedback. The encoders of two nearby UEs cooperate to extract the information of different channel paths under this framework, thereby reducing the overhead of repeated feedback of shared information. [12] considered the security issues of DL-based CSI feedback. [13] investigated interference and non-linear effects in practical scenarios.

Previous studies suggested that the architecture of the CSI feedback network is roughly divided into the feedforward neural network and the convolutional neural network (CNN). The latter exhibits more satisfactory performance than the former, but with extremely high computational complexity. We use a convolutional autoencoder in the present study to develop the network structure and complete the CSI feedback to preserve the spatial information of 2D signals effectively. The implementation of the CSI feedback network on the mobile terminal is challenging because this network, which is based on the CNN architecture, has excessive numbers of parameters and computational complexity. Thus, we adopt a lightweight structure based on the developed network to build a new one to reduce the complexity and number of parameters. The contributions of this study are summarized as follows.

We propose a DL-based CSI feedback structure called ConvCsiNet for FDD MIMO systems. This structure uses convolutional layers to extract features, while the mean-pooling and upsampling layers are used to compress and expand the matrix dimension multiples. Considering the application of mobile devices in the field of modern communications, we also propose a DL-based CSI feedback lightweight structure called ShuffleCsiNet. This structure can acquire accurate downlink CSI while consuming low memory space and core computing power. The experimental results show that ConvCsiNet is superior to CsiNet and other CsiNet-related works considering reconstruction performance. Moreover, ConvCsiNet improves the performance of the network at the cost of complexity. Compared with ConvCsiNet, the parameter number and algorithm complexity of ShuffleCsiNet are low, while the reconstruction performance is only slightly degraded.
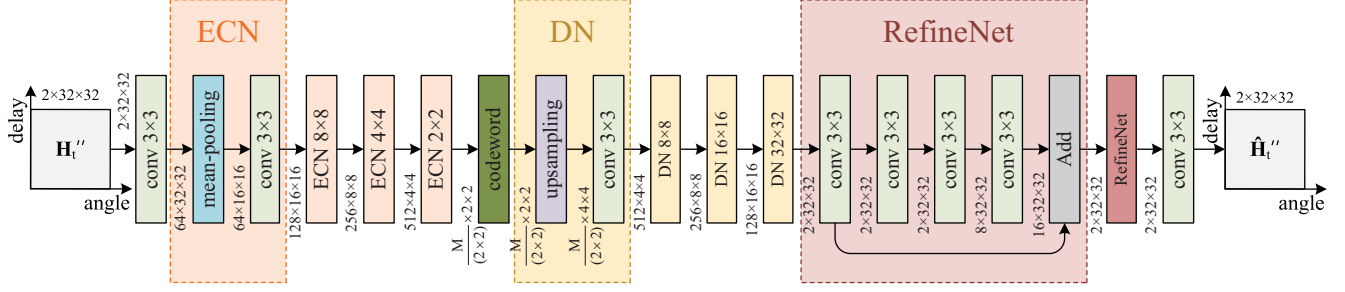
Fig. 1. Architecture of ConvCsiNet: An encoder structure constructed with convolutional layers and encoded convolution network (ECN) units, which include mean-pooling and convolutional layers; a decoder structure with convolutional layers, two series-connected RefineNet units, and DN units, which include upsampling and convolutional layers (**conv**: convolutional layer)

II. SYSTEM MODEL

This study considers a simple massive single-cell FDD downlink MIMO system with a BS and a UE. The BS is configured with a uniform linear antenna array, the number of transmitting antennas is $N_t \gg 1$, and the UE uses a single receive antenna. An OFDM system, which transmits information on N_c orthogonal subcarriers, is also considered. In a time-varying channel, the signal on the n $(n = 1, 2, \ldots, N_c)$ subcarrier received by the UE can be expressed as follows:

$$y_n = \mathbf{h}_n^T \mathbf{v}_n x_n + z_n, \tag{1}$$

where $\mathbf{h}_n \in \mathbb{C}^{N_t \times 1}$, $\mathbf{v}_n \in \mathbb{C}^{N_t \times 1}$, $x_n \in \mathbb{C}$, and $z_n \in \mathbb{C}$ respectively denote the instantaneous channel vector in the frequency domain, precoding vector, data symbol transmitted in the downlink, and additive Gaussian white noise. The channel vector on all subcarriers is the instantaneous downlink CSI matrix necessary for the UE to facilitate feedback. At this time, the CSI matrix can be denoted as $\mathbf{H} = [\mathbf{h}_1, \ldots, \mathbf{h}_{N_c}]^T \in \mathbb{C}^{N_c \times N_t}$. After receiving the CSI matrix $\mathbf{H}$ fed back by the UE, the BS can design the $\mathbf{v}_n$ precoding vector to reduce the interference among users. In actual situations, the UE must continuously estimate and feedback the instantaneous CSI to the BS to enable tracking of changes in the time-varying channel and adjust the corresponding precoding matrix. A total of $N_t N_c$ complex numbers must be transmitted for a complete CSI matrix due to a large number of antennas. These data occupy a substantial amount of feedback resources, which is undesirable for massive MIMO FDD systems in practical situations. Using a 2D discrete Fourier transform (2D-DFT), the CSI matrix $\mathbf{H}$ can be converted into an angular-delay domain matrix denoted as [4]

$$\mathbf{H}' = \mathbf{F_d} \mathbf{H} \mathbf{F_a}, \tag{2}$$

where $\mathbf{F_d} \in \mathbb{C}^{N_c \times N_c}$ and $\mathbf{F_a} \in \mathbb{C}^{N_t \times N_t}$ are two DFT matrices. As proven in [4], the CSI matrix $\mathbf{H}'$ is sparse in the angular-delay domain when performing DFT on the spatial domain channel vectors (i.e., row vectors of $\mathbf{H}$) if the number of transmit antennas $N_t \to +\infty$ is excessively large [14]. Only the first N_t' row contains non-zero values for the angular-delay domain channel matrix $\mathbf{H}'$. Therefore, we only keep the first N_t' rows of $\mathbf{H}'$ and obtain $\mathbf{H}'' \in \mathbb{C}^{N_c' \times N_t}$. The total number of parameters for the feedback is then reduced to $N = N_c' N_t$.

We mainly consider the autoencoder network for the down-link CSI feedback in this study. The channel matrix $\mathbf{H}''$ is transmitted to the autoencoder network, and the encoder compresses this matrix into codeword $\mathbf{s}$ according to a given compression ratio. After $\mathbf{s}$ is fed back to the BS, the decoder reconstructs $\mathbf{s}$ to $\mathbf{H}''$.

III. TWO PROPOSED DL-BASED CSI NETWORKS

We develop the DL-based CSI feedback network in this section using a convolutional autoencoder. On this basis, we then propose a new lightweight structured CSI feedback network. The specific model structure and parameter details are presented in the following subsections.

A. CSI feedback neural network architecture based on convolutional autoencoder

The CsiNet in [4] demonstrates superior performance to traditional CS algorithms considering CSI sensing and recovery. However, this algorithm uses the fully connected layer to change the dimensions. Moreover, CsiNet employs a fully connected autoencoder structure and a fully connected layer to generate codewords. Thus, this network cannot effectively retain the characteristics of 2D image signals. We propose a CSI feedback network based on a convolutional autoencoder called ConvCsiNet to address this issue. The neural network of ConvCsiNet uses convolutional layers instead of fully connected ones to extract features. In theory, the network trained under this architecture can satisfactorily extract the features of 2D image signals. Thus, ConvCsiNet is conducive to improving the reconstruction performance.

The architecture of ConvCsiNet shown in Fig. 1 is divided into the following two modules: encoder and decoder. The encoder includes a convolutional layer and four ECN units. The first convolutional layer contains 64 convolutional kernels with a size of 3×3 and a sliding step size of 1. The padding methods of this layer is set to "same" to generate a 64-channel feature map with the same size as the input CSI matrix to complete the initial feature extraction. The ECN is used to complete the dimensionality reduction and further feature extraction. Each ECN unit comprises average pooling and convolutional layers. The average pooling layer maintains the number of feature map channels and completes the downsampling which doubles the length and width of the feature map. The convolutional layers in the ECN units use the same features as those of the first convolutional layer. The

Fig. 2. (a) Encoder structure of ShuffleCsiNet constructed with convolutional and mean-pooling layers and SN units, which include convolutional, depthwise convolutional, concat, and channel shuffle layers. (b) Schematic of the channel shuffle layers (**conv**: convolutional layer; **dwconv**: depthwise convolutional layer)

convolutional layers of all other ECN units, except those of the last unit, maintain the feature map length and width, and the number of channels is doubled to further extract features. The number of channels of the convolutional layer in the last ECN unit is adjusted in accordance with the specific compression ratio. Each convolutional layer is added to the batch normalization (BN) layer and uses the Leaky ReLU activation function. The M-dimensional codeword $\mathbf{s}$ outputted by the encoder is presented in the form of a feature map with a size of 2×2 and the number of channels is $M/(2 \times 2)$.

The decoder part comprises four deconvolution network (DN) units, two RefineNet units, and a convolutional layer. The DN unit is used to restore codeword $\mathbf{s}$ to the original channel dimension and complete the preliminary reconstruction. Each DN unit comprises a bilinear up-sampling layer and a convolutional layer. The upsampling layer maintains the same number of feature map channels and completes the upsampling which doubles the length and width of the feature map. The convolutional layers in the DN units use a 3×3 convolutional kernel with a sliding step of 1 and the padding methods is set to "same". The convolutional layer of the first DN unit increases the number of feature map channels to 512 layers. By contrast, those of the second and third DN units maintain the feature map length and width, and the number of channels is doubled. The convolutional layer of the last DN unit generates a two-channel feature map, that is, two real number matrices, as the initial estimates of the real and imaginary parts of the original CSI matrix. These estimates then enter two similar RefineNet units to further improve the reconstruction quality. The last convolutional layer uses a 3×3 convolutional kernel with a sliding step size of 1 and the same padding and a Sigmoid activation function to normalize the output elements to the [0,1] interval. The remaining convolutional layers are added to the BN layer, and the Leaky ReLU activation function is used.

B. CSI feedback neural network architecture based on lightweight structure

The ConvCsiNet aims to optimize the CSI feedback network theoretically despite its complexity. Today, mobile terminals and embedded devices are widely used in the field of modern communications. The parameters and calculations of the ConvCsiNet network model are excessively large. Large memory space and core computing power are necessary to achieve an accurate downlink CSI. Therefore, the lightweight CSI feedback network model is important to ensure a sufficiently accurate training effect. The convolutional layer in ConvCsiNet has numerous parameters and computational complexity. Therefore, we propose a new autoencoder-based CSI feedback network called ShuffleCsiNet. This network replaces the average pooling and the convolutional layers with a lightweight downsampling structure.

The encoder and decoder parts of the CSI feedback network are located at the UE and BS, respectively. The mobile end has higher parameter requirements than those of the BS. Therefore, ShuffleCsiNet focuses on optimizing the structure of the encoder part, and the decoder part is the same as that in ConvCsiNet.

Fig. 2(a) shows that the encoder part of ShuffleCsiNet includes a convolutional layer, three shuffle network (SN) units, and a pooling structure. The first convolutional layer contains 64 convolutional kernels with a size of 3×3 and a sliding step size of 1. The padding methods of this layer is set to "same" to generate a 64-channel feature map with the same size as the input CSI matrix to complete the initial feature extraction. The SN is used to complete the dimensionality reduction and further feature extraction. Each SN unit divides the input into the following two parts: one part passes through a 3×3 depthwise convolutional layer and a 1×1 convolutional layer; the other part successively passes through a 1×1 convolutional layer, a 3×3 depthwise convolutional layer, and another 1×1 convolutional layer. The sliding step of the depthwise convolutional layer is two. The depthwise and the next convolutional layers with respective sizes of 3×3 and 1×1 can form a depthwise separable convolutional layer. This formed layer uses a 1×1 convolution kernel for the linear combination on the channel correlation and then performs convolution on the obtained feature map, which can jointly map the correlation of all dimensions. Moreover, this structure can achieve almost the same effect as the traditional convolution structure but with considerably reduced parameters. Afterward, the two branches pass the concat layer, which halves the size

of the feature space and doubles the number of channels to achieve feature reuse. Finally, channel shuffle can ensure information exchange between the two branches.

The SN unit can double the number of channels while maintaining the length and width of the feature map to further extract features. The pooling structure, which can also complete the downsampling function, comprises average pooling convolutional layers with a size of 3×3. Each depth separable convolutional layer is added to the BN layer, and each convolutional layer is added to the BN layer and uses the Leaky ReLU activation function. The M-dimensional codeword $\mathbf{s}$ outputted by the encoder is presented in the form of a feature map with a size of 2×2, and the number of channels is M/(2×2).

The structure of the channel shuffle is shown in Fig. 2(b). The channel shuffle can reduce element-level operations and increase the information flow after its placement in the concat layer [15]. The most important parameter in the channel shuffle structure is the feature map, that is, the number of shuffle groups. A wide feature map can obtain improved performance benefits within the range allowed by the model. After several experiments, we set the number of shuffle groups to 8 to obtain satisfactory reconstruction performance.

Let the transformation formula and all parameters of the entire network be $f(\cdot)$ and $\Theta = \{\Theta_{\mathsf{en}}, \Theta_{\mathsf{de}}\}$, respectively, where the parameters include encoder and decoder parameters. The CSI matrix recovered from the CSI feedback network model proposed in this study can then be expressed follows:

$$\hat{\mathbf{H}}'' = f(\mathbf{H}''; \Theta) \triangleq f_{\mathsf{de}}(f_{\mathsf{en}}(\mathbf{H}''; \Theta_{\mathsf{en}}); \Theta_{\mathsf{de}}). \quad (3)$$

We use the adaptive moment estimation (ADAM) algorithm to update the parameter set of the network. This algorithm is different from the traditional gradient descent algorithm, which uses a fixed learning rate. The ADAM algorithm can adaptively update the learning rate through training. The loss function of the network is represented by the mean squared error (MSE). Therefore, the prediction loss of the model is defined as follows:

$$L(\Theta) = \frac{1}{M} \sum_{m=1}^{M} \sum_{t=1}^{T} \|f(\mathbf{H}''; \Theta) - \mathbf{H}''\|_2^2. \quad (4)$$

where M is the total number of samples in the training set and $\| \cdot \|_2$ is the Euclidean norm.

IV. NUMERICAL RESULTS AND ANALYSIS

We compare and analyze the experimental data of ConvCsiNet and ShuffleCsiNet, including their reconstruction performance, parameter amount and FLOPs.

We use the COST 2100 model [16] to obtain the training and testing data and generate channel matrices for two environments. The two environments include indoor cellular and outdoor rural environments in the 5.3 GHz and 300 MHz bands, respectively. We set the bandwidth of the MIMO system to 20 MHz, and the number of subcarriers is $N_c = 256$. A uniform linear array with $N_t = 32$ antennas is used at the BS. The experiments in this study will be performed with CSI compression ratios (CR) of 1/16 and 1/32. The training, validation, and test sets used for offline training contain 75,000, 12,500, and 12,500 samples, respectively.

The entire training process is performed in the Keras framework. We utilize the ADAM optimizer using MSE loss to configure the training model. The parameters in ADAM are set as $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 1 \times 10^{-8}$. This study also uses a small batch training scheme with a batch size of 200, a training round set of 1000, and a learning rate of 0.001. CsiNet, ConvCsiNet, and ShuffleCsiNet are trained and tested on NVIDIA Tesla V100.

A. CSI reconstruction performance of the networks

The normalized MSE (NMSE) is used to evaluate the reconstruction performance and can be defined as follows:

$$\mathrm{NMSE} = \mathbb{E}\left\{\frac{1}{T} \sum_{t=1}^{T} \|\mathbf{H}'' - \hat{\mathbf{H}}''\|_2^2 / \|\mathbf{H}''\|_2^2\right\}. \quad (5)$$

The following cosine similarity is also calculated to facilitate comparison with the CsiNet.

$$\rho = \mathbb{E}\left\{\frac{1}{T} \frac{1}{N_c} \sum_{t=1}^{T} \sum_{n=1}^{N_c} \frac{|\hat{\mathbf{h}}_n^H \mathbf{h}_n|}{\|\hat{\mathbf{h}}_n\|_2 \|\mathbf{h}_n\|_2}\right\}, \quad (6)$$

where $\hat{\mathbf{h}}_n$ denotes the reconstructed channel vector of the nth subcarrier at time t, and ρ measures the quality of the beamforming vector when the vector is set as $\mathbf{v}_n = \hat{\mathbf{h}}_n / \|\hat{\mathbf{h}}_n\|_2$ because the UE will achieve the equivalent channel $\hat{\mathbf{h}}_n^H \hat{\mathbf{h}}_n / \|\hat{\mathbf{h}}_n\|_2$.

TABLE I
NMSE PERFORMANCE OF CSI FEEDBACK NETWORK DESIGNS.

CR	Method	Indoor		Outdoor	
		NMSE	ρ	NMSE	ρ
1/16	CsiNet	-8.65	0.93	-4.51	0.79
	CRNet	-11.35	0.95	-5.44	0.80
	DS-NLCsiNet	-12.45	0.97	-5.28	0.82
	ConvCsiNet	**-13.79**	**0.98**	**-6.00**	**0.85**
	ShuffleCsiNet	-12.14	0.97	-5.00	0.82
1/32	CsiNet	-6.24	0.89	-2.81	0.67
	CRNet	-8.93	0.94	-3.51	0.71
	DS-NLCsiNet	-8.21	0.92	-3.34	0.71
	ConvCsiNet	**-10.10**	**0.95**	**-5.21**	**0.82**
	ShuffleCsiNet	-9.41	0.94	-3.50	0.74

Table I summarizes the performance comparison considering NMSE. We compare the two proposed CSI feedback networks with CsiNet, CRNet [17], and DS-NLCsiNet [18]. The table shows that the reconstruction performances of ConvCsiNet and ShuffleCsiNet are superior to that of CsiNet. Compared with the two other CSI feedback networks, ConvCsiNet has an advantage when the compression ratio is 1/32. Notably, ShuffleCsiNet can maintain satisfactory reconstruction performance when parameter quantity and algorithm complexity are substantially reduced. Such a phenomenon is due to the functional reuse in the SN structure of the concat layer and due to the channel shuffle structure, thereby increasing the robustness of the network.

B. Parameter numbers and FLOPs of the networks

The ShuffleCsiNet aims to obtain a light network structure and low algorithm complexity. Therefore, the corresponding reconstruction performance of ShuffleCsiNet is slightly lower than that of ConvCsiNet. This study considers the convolutional , depthwise convolutional , BN layer, and fully connected layers when estimating the parameters of the network model. The parameter numbers of the convolution layer can be expressed as follows:

$$N_{\text{conv}} = K^2 \times C_i \times C_o, \tag{7}$$

where K is the size of the convolution kernel, C_i is the number of input channels, and C_o is the number of output channels. The parameter numbers of the depthwise convolutional, BN, and fully connected layers can be respectively expressed as follows:

$$\begin{aligned} N_{\text{dwconv}} &= K^2 \times C_i, \\ N_{\text{bn}} &= 2 \times C_i, \\ N_{\text{fc}} &= C_i \times C_o. \end{aligned} \tag{8}$$

Conversely, the convolution, depthwise convolutional, pooling, and fully connected layers are considered when estimating the FLOPS of the network model. The FLOPs of the convolution layer can be expressed as follows:

$$O_{\text{conv}} = H_o \times W_o \times K^2 \times C_i \times C_o, \tag{9}$$

where H_o, W_o, C_i, C_o, and K denote the height of the output, width of the output, number of input channels, number of output channels, and size of the convolution kernel, respectively. The FLOPs of the depthwise convolutional, pooling, and fully connected layers can be respectively expressed as follows:

$$\begin{aligned} O_{\text{dwconv}} &= H_o \times W_o \times K^2 \times C_i, \\ O_{\text{pool}} &= C_i \times H_i \times W_i, \\ O_{\text{fc}} &= 2 \times C_i \times C_o. \end{aligned} \tag{10}$$

Table II shows the model parameter numbers and FLOPs of the two proposed CSI feedback network designs. The model parameters of ShuffleCsiNet are much lower than those of ConvCsiNet. When the CR is 1/32, the parameter amount of ShuffleCsiNet can be reduced to 21.05% of that of ConvCsiNet. FLOPs can be used to measure the algorithmic complexity of network models. The results show that the complexity of the ShuffleCsiNet algorithm is considerably much lower than that of ConvCsiNet.

TABLE II
PARAMETER NUMBER AND FLOPS OF CSI FEEDBACK NETWORK DESIGNS.

CR	Method	numbers	FLOPs
1/16	ConvCsiNet	1,697,144	58,515,456
	ShuffleCsiNet	**415,528**	**11,845,632**
1/32	ConvCsiNet	1,623,416	58,220,544
	ShuffleCsiNet	**341,800**	**11,550,720**

V. CONCLUSION

We developed CsiNet and proposed a CSI feedback network called ConvCsiNet in this study based on a convolutional autoencoder. It can effectively extract the features of 2D image signals and exhibits satisfactory reconstruction performance. Subsequently, we proposed a lightweight structured CSI feedback network called ShuffleCsiNet based on ConvCsiNet. The experiments show that the proposed ConvCsiNet displays satisfactory reconstruction performance and the ShuffleCsiNet can substantially save memory space and kernel computing power while ensuring satisfactory reconstruction performance. Both CSI feedback architectures exhibit potentials for practical deployment on realistic MIMO systems.

REFERENCES

[1] F. Boccardi, R. W. Heath, A. Lozano, T. L. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, 2014.

[2] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, "Deep learning for wireless physical layer: Opportunities and challenges," *China Commun.*, vol. 14, no. 11, pp. 92–111, 2017.

[3] M. Chen, J. Guo, X. Li, and S. Jin, "An overview of the CSI feedback based on deep learning for massive MIMO systems," *Chin J. Internet Things*, vol. 4, no. 1, pp. 33–44, 2020.

[4] C.-K. Wen, W.-T. Shih, and S. Jin, "Deep learning for massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, 2018.

[5] T. Wang, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based CSI feedback approach for time-varying massive MIMO channels," *IEEE Wireless Commun. Lett.*, vol. 8, no. 2, pp. 416–419, 2018.

[6] Z. Liu, L. Zhang, and Z. Ding, "Exploiting bi-directional channel reciprocity in deep learning for low rate massive MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 889–892, 2019.

[7] X. Li and H. Wu, "Spatio-temporal representation with deep neural recurrent network in MIMO CSI feedback," *IEEE Wireless Commun. Lett.*, 2020.

[8] Y. Jang, G. Kong, M. Jung, S. Choi, and I.-M. Kim, "Deep autoencoder based CSI feedback with feedback errors and feedback delay in FDD massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 833–836, 2019.

[9] J. Guo, C.-K. Wen, S. Jin, and G. Y. Li, "Convolutional neural network based multiple-rate compressive sensing for massive MIMO CSI feedback: Design, simulation, and analysis," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2020.

[10] Q. Yang, M. B. Mashhadi, and D. Gündüz, "Deep convolutional compression for massive MIMO CSI feedback," in *Proc. IEEE Int. Conf. Mach. Learn. Signal Process. (MLSP)*, 2019, pp. 1–6.

[11] J. Guo, X. Yang, C.-K. Wen, S. Jin, and G. Y. Li, "DL-based CSI feedback and cooperative recovery in massive MIMO," *arXiv preprint arXiv:2003.03303*, 2020.

[12] Q. Liu, J. Guo, C.-K. Wen, and S. Jin, "Adversarial attack on DL-based massive MIMO CSI feedback," *arXiv preprint arXiv:2002.09896*, 2020.

[13] H. Ye, F. Gao, J. Qian, H. Wang, and G. Y. Li, "Deep learning based denoise network for CSI feedback in FDD massive MIMO systems," *arXiv preprint arXiv:2004.07576*, 2020.

[14] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, 2014.

[15] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, 2018, pp. 6848–6856.

[16] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. De Doncker, "The cost 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, 2012.

[17] Z. Lu, J. Wang, and J. Song, "Multi-resolution CSI feedback with deep learning in massive MIMO system," in *IEEE Int. Conf. Commun.*, 2020.

[18] X. Yu, Y. Bai, H. Wu, and X. Li, "DS-NLCsiNet: Exploiting non-local neural networks for massive MIMO CSI feedback," *arXiv*, 2020.