Channel Estimation for mmWave Massive MIMO With Convolutional Blind Denoising Network

Yu Jin, Jiayi Zhang, *Member, IEEE*, Bo Ai, *Senior Member, IEEE*, and Xiaodan Zhang

***Abstract*— Channel estimation is one of the foremost challenges for realizing practical millimeter-wave (mmWave) massive multiple-input multiple-output (MIMO) systems. To circumvent this problem, deep convolutional neural networks (CNNs) have been recently employed to achieve impressive success. However, current deep CNNs based channel estimators are only suitable to a small range of signal-to-noise ratios (SNRs). Unlike the existing works, the modified convolutional blind denoising network (CBDNet) is proposed to improve the robustness for noisy channel by adopting noise level estimation subnetwork, non-blind denosing subnetwork, and asymmetric joint loss functions for blind channel estimation. Furthermore, the CBDNet can adjust the estimated noise level map to interactively reduce the noise in the channel matrix. Numerical results demonstrate that the proposed CBDNet-based channel estimator can outperform the traditional channel estimators, traditional compressive sensing techniques and deep CNNs in terms of the normalized mean squared error. In addition, the CBDNet can be used over a large range of SNRs, which hugely reduce the cost of offline training.**

***Index Terms*— CBDNet, channel estimation, massive MIMO, millimeter wave.**

I. Introduction

TO ADDRESS the high throughput expectations of the fifth generation (5G) network, millimeter wave (mmWave) massive multiple input multiple output (MIMO) has become the key physical-layer technology [1]. Among signal processing techniques, the channel estimation is an indispensable step in realizing mmWave massive MIMO systems [2], [3]. More specifically, the acquisition of explicit channel state information (CSI) is extremely important for fully achieving the gain from wide mmWave bandwidths and large antenna arrays [4].

However, the implementation of conventional channel estimation schemes, such as minimum mean-squared error (MMSE) estimator, requires long-length pilot sequences and large-scale channel coefficients, which is unfeasible for practical mmWave massive MIMO systems due to very complicated channel characteristics. In mmWave massive MIMO systems, the channel is typically sparse in both time and angular dimensions due to the lack of scattering. In this context, a plethora of compressed sensing (CS) based channel estimation schemes have been introduced to fully leverage the inherent sparsity of mmWave massive MIMO channels. For example, the generalized approximate message passing (GAMP) based channel estimation was proposed in [5] especially for low SNRs and/or limited observations. By leveraging the sparse characteristic of mmWave channels, an off-grid sparse Bayesian learning algorithm was proposed for the channel estimation with a high estimation accuracy [6]. However, the state-of-the-art CS algorithms have potential limitations, such as high computational complexity because of the non-linear optimization. Moreover, as the mmWave channel sparsity patterns are often unknown, it is not efficient and reliable enough to realize accurate channel estimation by CS techniques.

Recently, the deep learning (DL) based framework has been successfully incorporated into the channel estimation for mmWave massive MIMO systems. DL is an effective tool to deal with the issue of complex nonlinear channel reconstruction. The authors in [7] considered the mmWave channel matrix as an 2D-image and leveraged a learned denoising-based approximate message passing (LDAMP) network to recover the channel. By exploiting both the spatial and frequency correlation, a spatial-frequency CNN (SF-CNN) based channel estimation for mmWave massive MIMO systems has been employed to achieve better performance [8]. Furthermore, a deep neural network (DNN) was proposed to realize the super-resolution channel estimation [9]. However, these DL-based channel estimation methods are restricted to a small range of noise level, which limits their practicality and applicability.

The convolutional blind denoising network (CBDNet) was developed to boost the blind denoising performance for real-world noisy images [10]. Motivated by this, we proposed the CBDNet-based channel estimation for mmWave massive MIMO systems. We model the channel matrix as an image. Then, the CBDNet is applied to exploit the sparsity characteristic and recover the channel matrix. Furthermore, an analytical framework on the normalized mean squared error (NMSE) performance of CBDNet is provided. From the simulation results, the CBDNet outperforms competing DL-based algorithms and can achieve notably performance gain with a wide range of SNRs and fast convergence.

Three significant improvement has been made: First, the hybrid neural network and nonlinear continuous output are used to improve the estimation accuracy and enlarge the SNR range; Second, the n-layer structure is utilized to improve the capacity of the subnetwork; Finally, the continuous nonlinear joint loss function is used to enlarge the SNR range and achieve fast convergence.

Manuscript received September 12, 2019; revised October 13, 2019; accepted November 4, 2019. Date of publication November 11, 2019; date of current version January 8, 2020. This work was supported by the Fundamental Research Funds for the Central Universities (2019JBM016). The associate editor coordinating the review of this letter and approving it for publication was J. Choi. *(Corresponding authors: Jiayi Zhang; Xiaodan Zhang.)*

Y. Jin and J. Zhang are with the School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: jiayizhang@bjtu.edu.cn).

B. Ai is with the State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing 100044, China, and also with the School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China.

X. Zhang is with the School of Management, Shenzhen Institute of Information Technology, Shenzhen 518172, China (e-mail: zhangxd@sziit.edu.cn).

Digital Object Identifier 10.1109/LCOMM.2019.2952845

II. System Model

Let us consider a typical mmWave massive MIMO system with arbitrary array geometry. N_r and N_t denote the number of antennas at receiver and transmitter, respectively. The uplink received signal $\mathbf{y}_m \in \mathbb{C}^{N_r \times N_t}$ is given by

$$\mathbf{y}_m = \mathbf{H}_k \mathbf{s} + \mathbf{n}, \tag{1}$$

where $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma_n^2 \mathbf{I}\right)$ is denoted as the Gaussian noise vector. Without loss of generality, we use the pilot vector $\left\|\mathbf{s}\right\|^2 = 1$ for channel estimation. According to [10], the proposed method can also be used in hybrid mmWave massive MIMO systems with non-Gaussian noise.

The channel between the BS and the kth UE is given by

$$\mathbf{H}_k = \sum_{l=1}^{L} z_l \boldsymbol{\alpha}_R\left(\phi_{R,l}^{\text{azi}}, \phi_{R,l}^{\text{ele}}\right) \boldsymbol{\alpha}_T^H\left(\phi_{T,l}^{\text{azi}}, \phi_{T,l}^{\text{ele}}\right), \tag{2}$$

where $L \ll \min\left(N_r, N_t\right)$ denotes the number of multipath and z_l donates the distance-dependent pathloss and shadowing. $\phi_{T,l}^{\text{ele}}$, $(\phi_{T,l}^{\text{azi}})$ and $\phi_{R,l}^{\text{ele}}$, $(\phi_{R,l}^{\text{azi}})$ denote the elevation (azimuth) angle of departure and angle of arrival of the lth path, respectively. $\boldsymbol{\alpha}_R\left(\phi_{T,l}^{\text{azi}}, \phi_{T,l}^{\text{ele}}\right)$ and $\boldsymbol{\alpha}_R\left(\phi_{R,l}^{\text{azi}}, \phi_{R,l}^{\text{ele}}\right)$ denote the steering vectors at the transmitter and the receiver, respectively.

These steering vectors depend on the array geometry. For the typical $N_1 \times N_2$ uniform planar arrays, $\boldsymbol{\alpha}\left(\phi_{T,l}^{\text{azi}}, \phi_{T,l}^{\text{ele}}\right)$ is given by [11]

$$\begin{aligned}\boldsymbol{\alpha}(\phi^{\text{azi}},\phi^{\text{ele}}) = & \left[1, e^{j2\pi d \sin\phi^{\text{azi}} \sin\phi^{\text{ele}}/\lambda}, \cdots, e^{j2\pi(N_1-1) d \sin\phi^{\text{azi}} \sin\phi^{\text{ele}}/\lambda}\right]^T \\ & \otimes \left[1, e^{j2\pi d \cos\phi^{\text{ele}}/\lambda}, \cdots, e^{j2\pi(N_2-1) d \cos\phi^{\text{ele}}/\lambda}\right]^T,\end{aligned} \tag{3}$$

where d is the antenna spacing, λ is the wavelength, and $\otimes$ denotes the Kronecker product.

Due to the large tightly-packet antenna array at the BS, antenna correlation increases in mmWave massive MIMO communication system. Thus, the channels are nearly sparse and the elements are correlated. This feature is highly similar to the 2D image. Consequently, the CBDNet, originated from the problem of image recovery, can be applied to exploit the channel estimation in the mmWave massive MIMO systems.

III. CBDNet Based Channel Estimation

The advanced CNNs have made major promotion in the field of Gaussian denoising. To further reduce the complexity of CNNs, blind denoising for real channel matrices include two phases, i.e., noise level estimation and non-blind denoising.

Fig. 1 illustrates the framework of CBDNet. To this end, the size of the received signal is $N_r \times N_t$. As the real and imaginary parts of the channel matrix can be seen as two independent channel matrices, we first combine them into a larger matrix of size $N_r \times 2\ N_t$. DNN$_E$ and DNN$_D$ denote the noise level estimation subnetwork and the non-blind denosing subnetwork, respectively. Furthermore, DNN$_E$ takes a noisy channel matrix $\mathbf{Y}$ to generate the estimated noise map $\mathbf{M}$ by training $\mathbf{W_E}$, which are network parameters of DNN$_E$. We have $\mathbf{M} = \mathcal{F}_E(\mathbf{Y}, \mathbf{W_E})$, where $\mathbf{M}$ has the same size of $\mathbf{Y}$.

In previous works, the CNN learns the structure of the noise and the noise level map does not have to be a part of the input. However, in the following asymmetric learning, it is necessary to estimate the noise level by improving the loss function. Thus, we can use $\mathbf{M}$ as part of the DNN$_D$ input. Then, the DNN$_D$ takes both $\mathbf{Y}$ and $\mathbf{M}$ as input to obtain the estimated channel $\mathbf{H} = \mathcal{F}_{\mathcal{D}}(\mathbf{Y}, \mathbf{M}, \mathbf{W_D})$, where $\mathbf{W_D}$ denotes the network parameters of DNN$_D$.

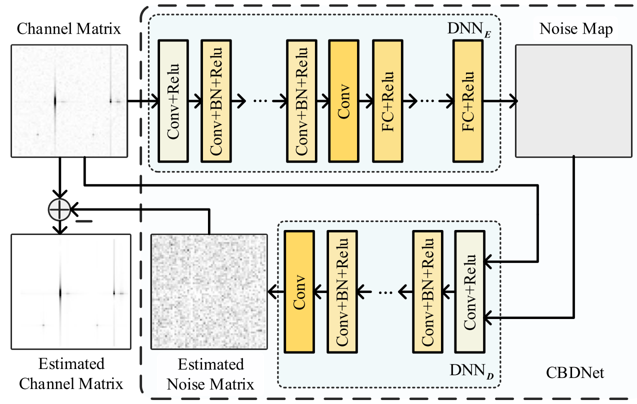

Fig. 1. CBDNet channel estimator for mmWave massive MIMO systems.

A. Noise Level Estimation Subnetwork

As shown in Fig. 1, DNN$_E$ consists of common fully connected (FC) layers and plain five fully convolution (Conv) layers. In each Conv layer, the number of feature channels is set as 32, and the filter size is 3×3. In the four middle FC layers, the number of connection points is 2000, 200, 50, 1, respectively. The ReLU nonlinearity is utilized after each middle Conv and FC layers.

In previous CNNs, DNN$_E$ trained the estimator model parameters σ using AWGN channel matrix with a fixed SNR level. But with fixed σ, the trained model lacks of flexibility in applying directly to channel matrices for other SNR levels. Although accurate channel estimation can work when the SNR level is within the preset training range, e.g., [17, 23] dB for FFDNet, the dynamic range is still very limited. In addition, all existing methods based on discriminative learning which lacks flexibility to handle spatial variant noise. For a space-invariant AWGN with a noise level of σ, M is a uniform map where all elements are σ.

B. Non-Blind Denosing Subnetwork

Different from DNN$_E$, the non-blind denoising subnetwork DNN$_D$ adopts an n-layer architecture which takes both $\mathbf{Y}$ and $\mathbf{M}$ as input to give a prediction $\mathbf{H}$ on the noise-free channel matrix. The residual learning is utilized by firstly learning the residual mapping $\mathcal{R}(\mathbf{Y}, \mathbf{M}, \mathbf{W_D})$ and then predicting $\mathbf{H} = \mathbf{Y} + \mathcal{R}(\mathbf{Y}, \mathbf{M}, \mathbf{W_D})$. The n-layer architecture of DNN$_E$ is given in Fig. 1, where symmetric skip connections, stridden convolutions and transpose convolutions are introduced for exploiting multi-scale information as well as enlarging receptive field. All the filter size is set 3×3, and the ReLU nonlinearity is employed after every Conv layer except the last one.

C. Asymmetric Learning

FFDNet is a non-blind CNN denoiser, which can obtain satisfying results on most channel matrices by manually setting proper or relatively higher noise level [12]. When the real input noise level is higher than the training SNR range, however, the result of FFDNet only contains the perceptible noise. Interestingly, when the input noise level is lower than the training SNR range, FFDNet can still achieve satisfactory results by gradually eliminating some low-contrast structures and increasing the input noise level [13]. Therefore, the non-blind CNN denoiser FFDNet is sensitive to the error of the high noise level. In mmWave massive MIMO systems, the channel estimation is often performed when the noise level is low and the error is robust. With such an attribute, FFDNet can be used for channel estimation by accepting relatively low input noise levels. However, FFDNet cannot suppress under-estimation error of noise level for high SNRs.

Thus, we use the asymmetric sensitivity of the non-blind denoiser, that is, different losses are used for different noise levels to improve the robustness of the CBDNet. To this end, the asymmetric loss is then introduced into the noise estimation subnetwork and use the combined reconstruction loss to train the CBDNet. Therefore, we present an asymmetric loss on the noise estimation to avoid the occurrence of under-estimation error on the noise level map. Given the estimated noise level $\sigma(\mathbf{Y})$ and the truth $\sigma(\mathbf{Y_i})$, more penalty is incorporated into their MSE when $\sigma(\mathbf{Y}) < \sigma(\mathbf{Y_i})$. The asymmetric loss on the noise estimation subnetwork is given by $\mathcal{L}_{\text{rec}} = \frac{1}{\sigma}\|\hat{\mathbf{H}} - \mathbf{H}\|^2$.

More specifically, multiple filters are used in the proposed CBDNet to obtain multiple features. The idea is to capture the most important feature with the largest value. One filter can be expressed as one feature. A series of 3×3 convolution multi-filters layers constitute the following CNN. Among them, one layer is combined by *Convolution* ("Conv"), *Rectified Linear Units* ("ReLU") and *Batch Normalization* ("BN"). Note that $\mathbf{y}_m$ is obtained as the input of DNN_E. Furthermore, $\mathbf{y}_m$ and $\mathbf{M}$ is obtained as the input of DNN_D. At the output of the last convolution layer, we obtain the estimated noise matrix with the size of $N_r \times 2N_t$.

D. Computational Complexity Analysis

For mmWave massive MIMO systems, the bottleneck of channel estimation is the high computational complexity. In CBDNet, the computational complexity of training phase mainly depends on the matrix multiplication, including both forward and backward propagation. The computational complexity of the training phase in CBDNet is given by

$$\begin{aligned}\mathcal{O}\Bigg(&4N_rN_tK^2s_Dt_D\left(D_2 + D_{L_d-1} + \sum_{l=2}^{L_d-1}\mathrm{D}_{l-1}D_l\right)\\&+2s_Dt_D\left(\mathrm{T}_1\mathrm{n}_1 + \mathrm{T}_L\mathrm{n}_{L_{FC}-1} + \sum_{l=2}^{L_{FC}-1} n_{l-1}n_l\right)\\&+4N_rN_tK^2s_Et_E\left(E_2 + E_{L_e-1} + \sum_{l=2}^{L_e-1}E_{l-1}E_l\right)\Bigg), \quad (4)\end{aligned}$$

where s donates the size of mini-batch, t donates the number of iterations, K^2 donates the size of kernels. L_d and L_e denote the total number of layers for DNN_D and DNN_E, D_l and E_l denote the number of features for the lth layer of DNN_D and DNN_E, respectively. Follow similar reason, the complexity of training phase of DnCNN is

$$\mathcal{O}\left(4N_rN_tK^2s_{Dn}t_{Dn}\left(Dn_2 + Dn_{L_{Dn}-1} + \sum_{l=2}^{L_{Dn}-1} Dn_{l-1}Dn_l\right)\right).$$

And the complexity of training phase of FFDNet is

$$\mathcal{O}\left(4N_rN_tK^2s_ft_f\left(F_2 + F_{L_f-1} + \sum_{l=2}^{L_f-1}F_{l-1}F_l\right)\right).$$

IV. Simulation Results

In this section, we investigate the NMSE performance of the proposed CBDNet over a wide range of SNRs. The results of other state-of-the-art baseline algorithms are also presented, such as FFDNet and DnCNN. Note that NMSE is defined as

$$\text{NMSE} = \mathbb{E}\left(\left\|\widehat{\mathbf{H}} - \mathbf{H}\right\|^2 / \left\|\widehat{\mathbf{H}}\right\|^2\right). \quad (5)$$

We consider the mmWave massive MIMO system, where $L = 3$ and $d = \lambda/2$, $N_r = 64$, $N_t = 32$. The path gains are assumed Gaussian distributed, i.e., $\alpha_l \sim \mathcal{CN}(0, \sigma_\alpha^2)$. The training rate is set to be 0.001, and training process stops when the error keeps unchanged in five sequential epochs. For other hyper-parameters of Adam, we use their default settings. The mini-batch size is set as 100, the rotation and flip based data augmentation is also adopted during training. The training, validation, and testing sets contain 16000, 6000, and 8000 samples, respectively. Furthermore, we scale the data to the range of [0, 100] for training the convolutional network. The CBDNet network is trained using the stochastic gradient descent (SGD) and Adam optimizer.

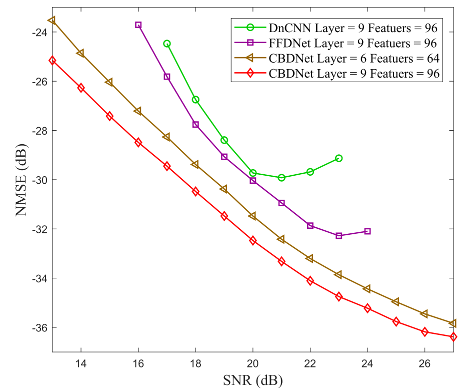

Fig. 2. NMSE performance of CBDNet, FFDNet and DnCNN with one single model ($N_r = 64$, $N_t = 32$).

Figure 2 compares the NMSE performance of CBDNet, FFDNet and DnCNN based channel estimator. DnCNN only fits one single noise level and lacks adaptability to other noise levels. FFDNet has a larger noise level than DnCNN. While CBDNet can deal with whole training SNR range. It also shows that FFDNet can achieve better NMSE performance when the range of SNR is [18, 22] dB, and DnCNN only match 20 dB. CBDNet shows a larger received field and achieves NMSE performance improvement with a smaller scale of CNN. Moreover, the performance of CBDNet can be further improved if using the same scale of CNN methods.

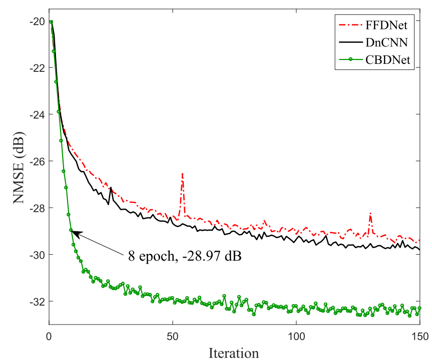

Fig. 3. Convergence of CBDNet, FFDNet and DnCNN ($N_r = 64$, $N_t = 32$, Layer = 10, Feature = 64, SNR = 20 dB).

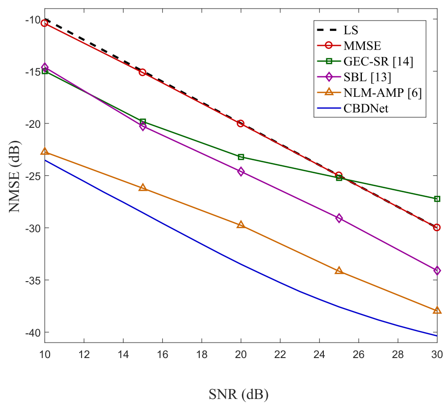

Fig. 4. NMSE performance comparison of CBDNet with traditional compressive sensing techniques (e.g., SBL, AMP, GEC-SR).

Figure 3 investigates the NMSE performance of different CNN models by using Sigmoid function as the activation function. Accordingly, when all CNNs have the same size, the CBDNet can achieve -28.97 dB NMSE within 8 epochs, while the DnCNN and FFDNet need to train more than 70 and 100 epochs. Compared with DnCNN or FFDNet, CBDNet achieves the best performance with faster convergence.

Figure 4 shows that the CBDNet is superior to the SBL algorithm [14]. For SNR = 15 and 25 dB, the NMSE of CBDNet is improved by 8.30 dB and 8.50 dB over the SBL algorithm, respectively. Furthermore, compared with NLM-AMP [7] and GEC-SR [15], the NMSE performance of CBDNet is also far superior for SNR in [10,30] dB. The computational complexity of VAMP is $\mathcal{O}\left(N^2R\right)$, $R = rank(\mathbf{A})$, where $\mathbf{A}$ is the measurement matrix. The computational complexity of GAMP is $\mathcal{O}\left(N^2\right)$. The computational complexity of GEC-SR is $\mathcal{O}\left(T_{max}N^3\alpha(\alpha+1)\right)$, where T_{max} is 35 and $\alpha = 0.7$ [15]. The computational complexity of SBL is $\mathcal{O}\left(T_{max}N^3\right)$, where $N = N_r^2$ and T_{max} is 150 [14].

Finally, we claim that the simulations are performed in PyCharm Community Edition on a computer with Intel(R) Core(TM) i5-7600K CPU @ 3.8GHz, 8 GB of RAM and an Nvidia GeForce GTX 1660Ti GPU. The average running times (in seconds) of CBDNet, FFDNet and DnCNN are 0.009364, 0.004073 and 0.004603, respectively. Note that the memory transfer time between GPU and CPU is also included.

V. Conclusion

In this letter, we present a CBDNet-based channel estimator for mmWave massive MIMO systems. For model learning, asymmetric loss is introduced to the noise estimation subnetwork to impose higher penalty on under-estimation error of noise level, and reconstruction loss is adopted to train the network. To improve the generalization ability of a large noise level, we incorporate both residual and asymmetric learning for loss function during training of the CBDNet. Numerical results verify that the proposed CBDNet can achieve better NMSE performance with lower computational complexity.

References

[1] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Multiple antenna technologies for beyond 5G," 2019, *arXiv:1910.00092*. [Online]. Available: https://arxiv.org/abs/1910.00092
[2] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, "On low-resolution ADCs in practical 5G millimeter-wave massive MIMO systems," *IEEE Commun. Mag.*, vol. 56, no. 7, pp. 205–211, Jul. 2018.
[3] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
[4] J. Zhang, L. Dai, Z. He, B. Ai, and O. A. Dobre, "Mixed-ADC/DAC multipair massive MIMO relaying systems: Performance analysis and power optimization," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 140–153, Jan. 2019.
[5] F. Bellili, F. Sohrabi, and W. Yu, "Generalized approximate message passing for massive MIMO mm Wave channel estimation with laplacian prior," *IEEE Trans. Commun.*, vol. 67, no. 5, pp. 3205–3219, May 2019.
[6] H. Tang, J. Wang, and L. He, "Off-grid sparse Bayesian learning-based channel estimation for Mm Wave massive MIMO uplink," *IEEE Wireless Commun. Lett.*, vol. 8, no. 1, pp. 45–48, Feb. 2019.
[7] H. He, C.-K. Wen, S. Jin, and G. Y. Li, "Deep learning-based channel estimation for beamspace mmWave massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 852–855, Oct. 2018.
[8] P. Dong, H. Zhang, G. Y. Li, I. S. Gaspar, and N. NaderiAlizadeh, "Deep CNN-based channel estimation for mm Wave massive MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 5, pp. 989–1000, Sep. 2019.
[9] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, "Deep learning for super-resolution channel estimation and doa estimation based massive MIMO system," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.
[10] S. Guo, Z. Yan, K. Zhang, W. Zuo, and L. Zhang, "Toward convolutional blind denoising of real photographs," in *Proc. IEEE CVPR*, Jun. 2019, pp. 1712–1722.
[11] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, Jr., "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
[12] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
[13] Y. Jin, J. Zhang, S. Jin, and B. Ai, "Channel estimation for cell-free mm Wave massive MIMO through deep learning," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10325–10329, Oct. 2019.
[14] A. Mishra, A. Rajoriya, A. K. Jagannatham, and G. Ascheid, "Sparse Bayesian learning-based channel estimation in millimeter wave hybrid MIMO systems," in *Proc. IEEE SPAWC*, Jul. 2017, pp. 1–5.
[15] H. He, C.-K. Wen, and S. Jin, "Generalized expectation consistent signal recovery for nonlinear measurements," in *Proc. IEEE ISIT*, Jun. 2017, pp. 2333–2337.