

# A Unified Deep Learning Method for CSI Feedback in Massive MIMO Systems



GAO Zhengguang<sup>1</sup>, LI Lun<sup>1</sup>, WU Hao<sup>1</sup>,

TU Xuezheng<sup>2</sup>, HAN Bingtao<sup>1</sup>

(1. State Key Laboratory of Mobile Network and Mobile Multimedia Technology, Shenzhen 518055, China;

2. The College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

DOI: 10.12142/ZTECOM.202204013

<https://kns.cnki.net/kcms/detail/34.1294.TN.20220426.0828.002.html>  
published online April 26, 2022

Manuscript received: 2021-11-15

**Abstract:** A unified deep learning (DL) based algorithm is proposed for channel state information (CSI) compression in massive multiple-input multiple-output (MIMO) systems. More importantly, the element filling strategy is investigated to address the problem of model redesigning and retraining for different antenna typologies in practical systems. The results show that the proposed DL-based algorithm achieves better performance than the enhanced Type II algorithm in Release 16 of 3GPP. The proposed element filling strategy enables one-time training of a unified model to compress and reconstruct different channel state matrices in a practical MIMO system.

**Keywords:** deep learning; channel state information; element filling strategy

**Citation** (IEEE Format): Z. G. Gao, L. Li, H. Wu, et al., "A unified deep learning method for CSI feedback in massive MIMO systems," *ZTE Communications*, vol. 20, no. 4, pp. 110 - 115, Dec. 2022. doi: 10.12142/ZTECOM.202204013.

## 1 Introduction

In 5G and Beyond networks, massive multiple-input multiple-output (MIMO) is considered one of the key enabling technologies to improve link capacity and energy efficiency for wireless communications<sup>[1-5]</sup>. To achieve these potential advantages, simultaneous channel state information (CSI) is required to optimize the precoding for massive MIMO systems. In frequency division duplex (FDD) MIMO systems of 4G, the downlink CSI obtained at the user equipment (UE) is sent to the base station, and vector quantization or codebook-based approaches are adopted as the compression algorithms for CSI to decrease feedback overhead<sup>[6]</sup>. However, the feedback overhead increases significantly in massive MIMO systems because the feedback quantity of the current methods increases linearly with the number of antennas. This challenge has inspired researchers to explore an effective algorithm to compress the CSI in massive MIMO systems. The technology of compressive sensing (CS) is exploited to address this issue<sup>[7-8]</sup>. Based on the uncorrelated sparse vector transformed from the correlated CSI matrix, the CS-based methods are expected to achieve an accurate performance for CSI compression. Several CS-based algorithms have been proposed in massive MIMO systems, such as the least absolute shrinkage and selection operator (LASSO)  $l_1$ -solver and approximate message passing (AMP)<sup>[9]</sup>. The advanced CS-based algorithms which include TV minimization by augmented Lagrangian and

alternating direction algorithms (TVAL3) as well as block-matching and 3D filtering (BM3D)-AMP have also been proposed to improve the accuracy<sup>[10-11]</sup>. Although CS-based methods have been investigated comprehensively, it still has inherent disadvantages<sup>[11]</sup>. Firstly, the effectiveness of CS-based methods relies on the assumption that CSI matrices are sparse in some bases whereas channels are not always sparse in practical systems. And the random projection of CS-based methods cannot extract useful information from the channel structure of MIMO systems, which has negative impacts on the algorithm performance. Finally, the decoding process always requires iterative solving, therefore the decompression of the CS-based method is sub-optimal and time-consuming. In a word, the CS-based method cannot achieve high performance for the reason that the CSI matrix is not sparse enough under the large compression ratio, and the slow reconstruction of the CS-based method makes it difficult to adapt many real-time scenarios in practical systems. Recently, deep learning has been explored in signal detection and network planning<sup>[12-13]</sup>. Motivated by the rapid progress of deep learning in computer vision (CV), in particular, the successful trial of image compression and reconstruction by the autoencoder, the researcher has explored DL-based algorithms for CSI compression comprehensively.

A novel residual neural network-based model called CsiNet is proposed in Ref. [14], which shows that the performance of CsiNet outperforms existing CS-based methods significantly,

especially for low compression regions. To exploit the temporal correlations of CSI, a long-short time memory (LSTM) architecture is combined with CsiNet as CsiNet-LSTM in Ref. [15], and it shows that considering the temporal correlations benefits the accuracy of CSI reconstruction. The CsiNet is modified and redesigned as CsiNet+ in Ref. [16], which improves the performance of CSI compression. More importantly, a novel quantization layer is introduced in the DL model for end-to-end training, which meets the practical requirement of CSI feedback in massive MIMO systems. Recently two novel DL-based models called CRNet and ACRNet have been proposed for better performance in Refs. [17 - 18]. In Ref. [17], multi-resolution CRBlocks are designed in CRNet, and the warm-up aided training scheduler is implemented to achieve better performance. The result proves the effectiveness of multi-resolution CRBlocks and the novel training scheduler. In Ref. [18], a novel model called ACRNet is proposed to provide the state-of-the-art performance with network aggregation and parametric rectified linear unit (PReLU) activation. Besides, the network binarization technique is implemented to ensure the high performance and small memory cost. Most of the above works have achieved the state-of-art performance previously and outperformed CS-based methods significantly in some regions, while some problems still exist in the following aspects. Most enlightening works focus on novel designs of the DL model to improve the performance of CSI compression, but ignore the quantization of information in bit-streams for transmission in practical MIMO systems. These models should consider the quantization layer and have end-to-end training and testing to evaluate the performance of practical systems. In addition, the proposed DL-based algorithms are evaluated in the COST2100 dataset<sup>[19]</sup>, which only includes data with the same shape. However, the shape of the CSI matrix changes with antenna arrangements in massive MIMO systems. Generally, misalignment problems exist between the previous model and the coming data with different shapes. To our knowledge, there are no published works discussing how to deal with this issue in practical systems. Motivated by this, a unified DL-based method for CSI feedback is proposed in massive MIMO Systems. The rest of this paper are organized as follows.

A novel DL-based network named ACRNet+ is proposed in Section 2. To improve the performance in the practical massive MIMO system, the advanced module of channel attention and spatial attention is used to enhance the ability of feature extraction. And a 3-bit uniform quantization layer is implemented in ACRNet+ for end-to-end training in Section 3, which minimizes the impact of the accuracy quantization of the model. In addition, an element filling strategy is proposed to address the incompatible problem of the trained model for the CSI matrix with a different size. Section 4 shows the experimental results that ACRNet+ provides better performance than the enhanced Type II (eType II) algorithm in the 901 dataset which focuses on outdoor scenarios with dual-polarized

antennas. The element filling strategy enables the trained model to compress the diversified CSI matrices accurately without further training. Section 5 concludes the paper.

## 2 System Model

A simple single-cell downlink is adopted in massive MIMO systems with  $N_t$  antennas at the base station (BS) and  $N_r$  antennas at UE, where  $N_t \geq 1$  and  $N_r = 1$  in this paper. The system considers orthogonal frequency division multiplexing (OFDM) with subcarrier  $N_c$ <sup>[20]</sup>, so the received signal  $y_n$  at the  $n$ -th subcarrier can be expressed as:

$$y_n = \mathbf{h}_n^H \mathbf{p}_n x_n + z_n, \quad (1)$$

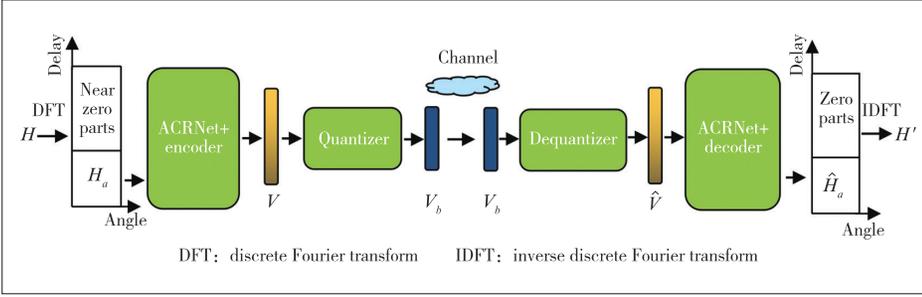
where  $\mathbf{h}_n \in \mathbb{C}^{N_t \times 1}$  is the downlink channel vector, and  $\mathbf{p}_n \in \mathbb{C}^{N_t \times 1}$  is the precoding vector based on CSI.  $x_n \in \mathbb{C}$  and  $z_n \in \mathbb{C}$  are the transmitted symbol and additive Gaussian noise for the  $n$ -th subcarrier. To obtain the advantage of channel gain in massive MIMO systems, the downlink CSI should be acquired at the BS to optimize the channel precoding. Since the downlink channel matrix  $\hat{\mathbf{H}} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_{N_c}]^H$  contains  $2 \times N_c \times N_t$  elements of the float type, the direct feedback of CSI for downlink is not feasible obviously. In order to reduce the overhead of CSI feedback, a 2D discrete Fourier transform (DFT) is used to make the matrix sparse in the angular delay domain as<sup>[21]</sup>

$$\mathbf{H} = \mathbf{F}_d \hat{\mathbf{H}} \mathbf{F}_a^H, \quad (2)$$

where  $\mathbf{F}_d$  and  $\mathbf{F}_a^H$  are DFT matrices with the shape of  $N_c \times N_c$  and  $N_t \times N_t$  respectively. The obtained matrix  $\mathbf{H}$  contains only the first  $N_a$  rows of useful elements whereas other rows of elements are near zero. Although the DFT can reduce the elements of the downlink channel matrix, the first  $N_a$  row of the transformed matrix called  $\mathbf{H}_a$  is still too large to be sent directly through the uplink channel. According to the inaccuracy of the large compression ratio and the slowing decoding procedure of CS-based methods, a DL-based encoder and decoder are considered for CSI compression at UE and reconstruction at the BS. An auto-encoder is a neural network-based model to reconstruct the raw data through self-supervised learning. It firstly builds the main features into a lower-dimensional representation of the input data, and then the decoder tries to reconstruct the data as similarly as possible<sup>[22]</sup>.

The overview of the DL-based model for uplink feedback is shown in Fig. 1. The CSI matrix  $\mathbf{H}_a$  in the angular-delay domain is obtained from the DFT of  $\hat{\mathbf{H}}$  and matrix truncation. The encoder of ACRNet+ compresses the input of  $\mathbf{H}_a$  into a one-dimensional vector  $\mathbf{v}$  with  $M$  elements. Therefore, the compressive ratio can be expressed as

$$\eta = \frac{M}{2N_a N_t}. \quad (3)$$



▲ Figure 1. Pipeline of ACRNet+ model for channel state information (CSI) compression

In practical systems, the bitstream is required for uplink feedback, so a 3-bit uniform quantizer is used to convert the float vector  $\mathbf{v}$  into a binary vector  $\mathbf{v}_b$ . The above process is abstracted as

$$\mathbf{H} = F_q(F_{en}(\mathbf{H}_a, \mathbf{W}_{en})), \quad (4)$$

where  $F_q$  denotes the function of quantization,  $F_{en}$  stands for the encoder of ACRNet+, and  $\mathbf{W}_{en}$  is the trainable parameter of the encoder. After uplink transmission, the bitstream  $\mathbf{v}_b$  is dequantized into the float vector, and then the float feature is reconstructed into  $\hat{\mathbf{H}}_a$ . The decoding process can be formulated as

$$\hat{\mathbf{H}}_a = F_{de}(F_{dq}(\mathbf{v}_b), \mathbf{W}_{de}), \quad (5)$$

where  $F_{dq}$  denotes the function of dequantization,  $F_{de}$  stands for the decoder of ACRNet+, and  $\mathbf{W}_{de}$  is the trainable parameter of the decoder. After the decoding process, the estimated channel matrix for the downlink can be obtained from the inverse DFT of  $\hat{\mathbf{H}}_a$ .

## 3 Description of ACRNet+

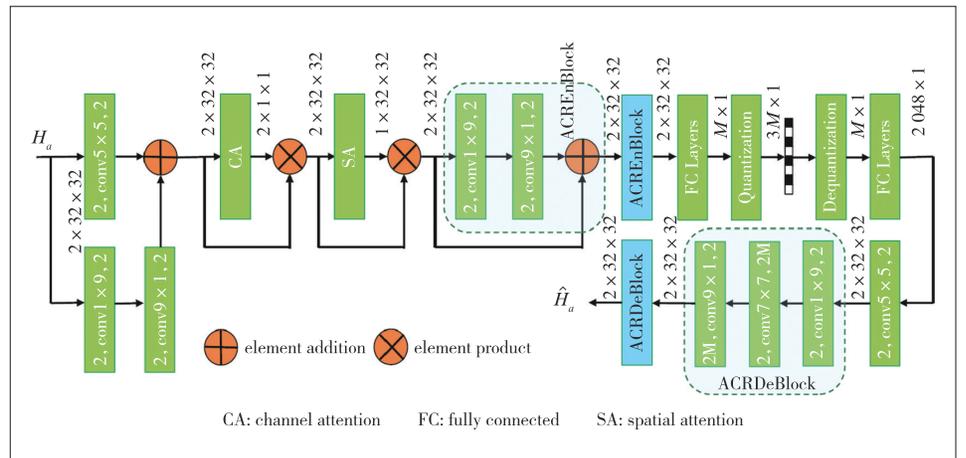
### 3.1 Design of ACRNet+

ACRNet is one of the most effective models in existing works, which provides the state-of-art performance for CSI compression in COST2100<sup>[18]</sup>. Based on the network aggregation technique, a novel feedback block that contains parallel convolutional groups is implemented to extract relatively independent features<sup>[23–24]</sup>. In addition, a learnable activation function of the parametric ReLU and an advanced training scheme are used to boost the performance of the model.

To inherit the advantages of ACRNet, ACREnBlock and ACRDeBlock are introduced in the proposed ACRNet+. In addition,

channel attention (CA) and spatial attention (SA) are implemented to extract the useful features accurately from the CSI matrices, especially for outdoor scenarios with the randomness of multi-path fading and unpredicted interferences<sup>[25]</sup>. Inspired by the design of multi-resolution networks for CSI feedback tasks in previous works<sup>[17]</sup>, two independent channels are built to provide more feature granularity. After the combination of the features from these two channels, CA and SA are used to make further extraction of the combined features. Different from the design of ACRNet, a quantized layer is added to convert the compressed data into bitstreams for feedback transmission on the uplink. At the BS, the bitstream is dequantized into the float features, and then the compressed vector from the dequantized layer is reconstructed into  $\hat{\mathbf{H}}_a$  through one fully connected network and two ACRDeBlocks.

The detailed architecture of ACRNet+ is shown in Fig. 2. The image input  $\mathbf{H}_a$  with the shape  $2 \times H_a \times H_t$  is fed into two independent channels. The first channel includes a  $5 \times 5$  convolution layer, and the second channel is made up of two convolution layers with a  $1 \times 9$  kernel and a  $9 \times 1$  kernel respectively. Then the outputs of the two channels are combined through element addition. To boost the representation power of networks, a convolutional block attention module (CBAM) is implemented to focus on important features and suppress unnecessary ones along two separate dimensions<sup>[25]</sup>. After the adaptive feature refinement by CBAM, the output of CBAM passes through the two ACREnBlocks which include a  $1 \times 9$  kernel and a  $9 \times 1$  kernel for further feature processing, and then the flattened feature from ACREnBlock is condensed into the vector with  $M$  elements. To meet the practical requirement for transmission of wireless communication, a 3-bit uniform quantizer is imple-



▲ Figure 2. Proposed ACRNet+ architecture

mented for quantization. The quantizer is embedded in neural networks as the quantization layer for end-to-end training, which can minimize the negative impacts on the model accuracy. At the BS, the bitstream is firstly recovered to the compressed vector through dequantization. Then the vector passes through a fully connected layer for element augmentation, the output is reshaped in the same dimension as the input. Finally, the reshaped feature passes one convolution layer with a  $5 \times 5$  kernel and two ACRDeblocks which include a  $1 \times 9$  kernel, a  $7 \times 7$  kernel and a  $9 \times 1$  kernel for the final reconstruction.

### 3.2 Generalization Scheme for CSI Compression

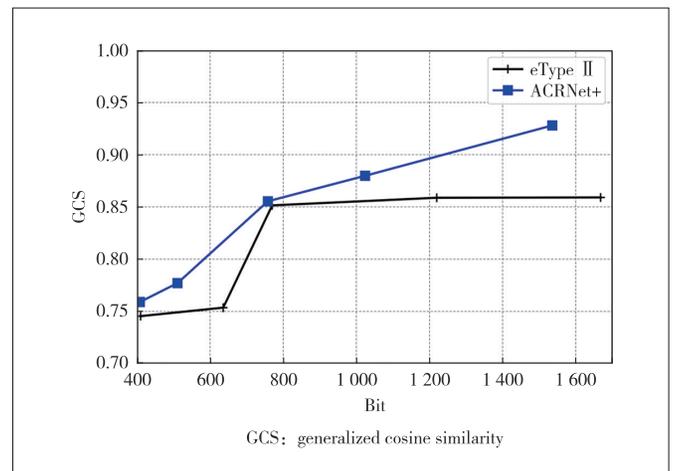
One of the main concerns for DL-based methods is the generalization problem. For data with the same shape, we can evaluate the model through the performance of cross validation and testing. DL-based methods for CSI compression have been proven effective in most previous works. However, the generalization of the DL-based model has never been discussed for data with different shapes. In practical MIMO systems, the shape of the downlink channel matrix changes with the arrangement of antennas. The misalignment problem happens in this scenario. The direct method is to get several models trained for different antenna typologies, but it requires huge computation resources for several well-trained models as well as the corresponding switching strategy. Therefore, the element filling strategy is proposed to support the training of a unified model for different scenarios covering several antenna combinations. At first, the dataset with the largest size is selected as the benchmark. Then the dataset with other shapes is filled by the constant elements with the same shape as the benchmark. Finally, a unified DL-based model is trained on the hybrid dataset covering the data from several antenna combinations in massive MIMO systems. Considering the fact that most elements in  $H_a$  are close to 0.5, the padding element is set to 0.5, which can decrease the disturbance of the padding elements for the compression and reconstruction of the actual elements.

In this paper, we consider three scenarios including  $4 \times 4$  MIMO,  $2 \times 4$  MIMO, and  $1 \times 8$  MIMO with dual-polarized antennas, and the corresponding datasets are represented as data1, data2 and data3. During the training process, the real and imaginary parts of the data are represented as the third dimension, and the shapes for these datasets are  $2 \times 32 \times 32$ ,  $2 \times 16 \times 32$ , and  $2 \times 16 \times 32$  respectively. According to the element filling strategy, we perform the padding operation to data2 and data3, which reshapes them as the size of  $2 \times 32 \times 32$ . After the same shape for these scenarios is obtained, a unified model is trained on the dataset which randomly samples from data1, data2 and data3. We compare the performance of the model trained on the hybrid dataset with the model trained separately on one of the three datasets.

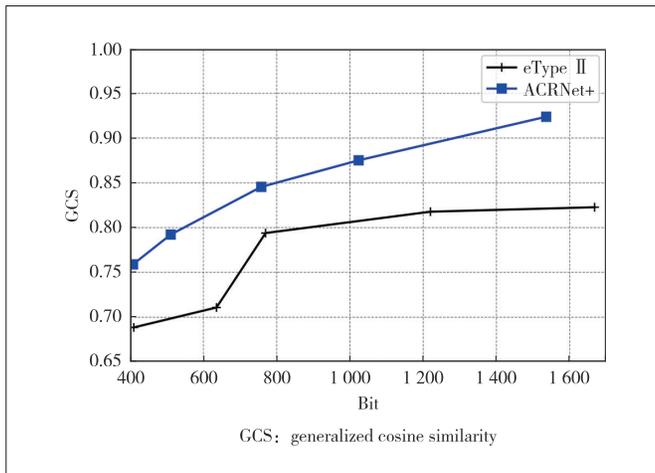
## 4 Experiment Results and Analysis

Most previous works are based on COST2100, which deploys a uniform linear array with 32 antennas at the BS and 1 024 subcarriers. In this paper, more complicated scenarios are considered to evaluate the effectiveness of the DL-based model. To evaluate the practical performance of ACRNet+, we first compare the performance of ACRNet+ with eType2 algorithm of Release 16 (R16) from 3GPP in two outdoor scenarios with dual-polarized antennas and two antenna topologies. The shape of the two datasets is  $2 \times 32 \times 32$ . In addition, a unified model is trained on three datasets with different shapes through the element filling strategy. The performance of ACRNet+ trained on the hybrid dataset is compared with ACRNet+ trained separately on one of these datasets.

In order to evaluate the proposed model in more practical scenarios, 901data is generated according to the 3GPP Long Term Evolution (LTE) structure, in which parameters of channel models are specified in TR38: 901<sup>[26]</sup>. We compare ACRNet+ and eType2 algorithms on 901data containing two datasets, in which  $16 \times 1$  MIMO and  $8 \times 2$  MIMO with dual-polarized antennas are used. The generalized cosine similarity (GCS) of these two methods is presented in Figs. 3 and 4. The figures show that ACRNet+ outperforms eType II in terms of GCS when consuming similar bits. For the scenario in Fig. 3, the GCS of eType II is 0.745 4 under 410 bits. The corresponding GCS of ACRNet+ is 0.758 9 under 408 bits. It should be noted that it is difficult to completely match the number of bits for the two algorithms on the  $x$ -axis. The number of bits for eType II is 410, 636, 770, 1 220, and 1 668, and the consumed bits for ACRNet+ are 408, 511, 768, 1 024, and 1 536. We can see that the GCS of ACRNet+ is better than that of eType II although the ACRNet+ requires fewer bits. For the scenario in Fig. 4, the advantage of ACRNet+ is more obvious. The closest performance for ACRNet+ and eType II occurs when the number of bits is about 770. The corresponding GCS for eType II is about 0.793 9 where the



▲ Figure 3. GCS of ACRNet+ and eType II for the scenario with  $1 \times 16$  multiple-input multiple-output (MIMO)



▲ Figure 4. GCS of ACRNet+ and eType II for the scenario with 2×8 multiple-input multiple-output (MIMO)

GCS of ACRNet+ under 768 bits is about 0.845 8. Finally, it can be seen from Fig. 3 that the GCSs of the algorithms are not linearly related to the consumed bits. The GCS cannot be significantly improved when the used bit is below some threshold, and the improvement is also unobvious when the consumed bit is enough for the eType II algorithm. These results provide a useful reference for the actual deployment of these algorithms.

One of the main concerns for DL-based CSI compression methods is the generalization of the model. When the topology of antennas changes, the coming data of the channel matrix are incompatible with the original DL-based model. The re-training of the model requires huge computation resources and training time. Therefore, an element filling strategy is proposed to unify the input shape. Then a unified model is trained on the datasets which cover several scenarios with different typologies. In the experiment, three datasets including 4×4 MIMO, 2×4 MIMO, and 1×8 MIMO with dual-polarized antennas are generated in outdoor scenarios, and the corresponding datasets are represented as data1, data2 and data3. The training, validation, and testing sets for each dataset contain 85 000, 5 000, and 10 000 samples. The shapes of these datasets are  $2 \times 32 \times 32$ ,  $2 \times 16 \times 32$ , and  $2 \times 16 \times 32$  respectively. At first, data2 and data3 are reshaped into  $2 \times 32 \times 32$ , where the padding element is 0.5 for all the places. We mix the reshaped datasets and randomly sample one-third of the mixed data as the training set. Therefore, the mixed training dataset also contains 85 000 samples. Finally, the performance of ACRNet+ trained on the mixed dataset is compared with the model trained on the original dataset to evaluate the effectiveness of the element filling strategy.

The results of ACRNet+ trained on the mixed dataset and the original dataset are presented in Table 1, in which ACRNetH represents the performance of the model trained on a mixed dataset and ACRNet+ stands for the model trained separately on the original dataset. We can see from the table

that the performance of ACRNetH which needs to be trained once is only slightly lower than ACRNet+ which needs retraining for data with different antenna topologies. For example, the GCS of ACRNet+ under 15 times compression ratio is 0.746 2, 0.875 4 and 0.844 9, and the performance of ACRNetH that only requires one-time training reaches a similar performance as 0.722 8, 0.850 9 and 0.810 5. The results show that the proposed element filling strategy enables the unified training of the model for the dataset which contains the samples with different sizes, and the corresponding performance reaches a similar performance trained separately.

▼ Table 1. Comparison between ACRNet+ and ACRNetH under the same compression ratio

Dataset	Method	1/4		1/8		1/15	
		NMSE	GCS	NMSE	GCS	NMSE	GCS
Data1	ACRNet+	-7.22	0.887 9	-5.10	0.813 4	-3.865	0.746 2
	ACRNetH	-6.59	0.870 0	-4.70	0.795 8	-3.44	0.722 8
Data2	ACRNet+	-14.10	0.976 5	-9.78	0.937 5	-6.61	0.875 4
	ACRNetH	-12.38	0.966 1	-8.07	0.910 5	-5.82	0.850 9
Data3	ACRNet+	-12.62	0.969 8	-8.001	0.911 4	-5.678	0.844 9
	ACRNetH	-9.84	0.943	-6.812	0.882 5	-4.795	0.810 5

GCS: generalized cosine similarity

NMSE: normalized mean squared error

One thing that should be noted is that the proposed model as well as the mentioned models in the paper addresses the CSI compression in low-speed scenarios within 3 km/h, while unknown and complex factors on the performance of the CSI feedback still exist in high-speed scenarios, which we plan to analyze in future works.

## 5 Conclusions

In this paper, a unified DL-based model called ACRNet+ has been proposed to compress the CSI in massive MIMO systems. The proposed model outperforms eType II algorithms in R16 of 3GPP in two outdoor scenarios. More importantly, the element filling strategy allows a unified training of the model on the dataset containing samples with different shapes, which enables the one-time training of DL-based model to address the CSI compression for different antenna typologies. The experimental results show that the performance of a unified model can reach a similar performance of the DL-based model trained separately for the dataset of a certain scenario.

## Acknowledgement

Our gratitude goes to WANG Yongcheng, YANG Xikun and LUO Qingkai for the technical advices and the linguistic assistance.

## References

- [1] LU L, LI G Y, SWINDLEHURST A L, et al. An overview of massive MIMO: benefits and challenges [J]. IEEE journal of selected topics in signal processing,

- 2014, 8(5): 742 – 758. DOI: 10.1109/JSTSP.2014.2317671
- [2] MARZETTA T L. Massive MIMO: an introduction [J]. Bell labs technical journal, 2015, 20: 11 – 22. DOI: 10.15325/BLTJ.2015.2407793
- [3] WU H Q. Ten reflections on 5G [J]. ZTE technology journal, 2020, 26(1): 2 – 4. DOI: 10.12142/ZTECOM.202001001
- [4] FANG M, DUAN X Y, HU L J. Challenges, innovations and perspectives towards 6G [J]. ZTE technology journal, 2020, 26(3): 61 – 70. DOI: 10.12142/ZTETJ.202003012
- [5] WANG X Y. 5G: striving for sustainable growth amid expectations [J]. ZTE technology journal, 2020, 26(1): 64 – 66. DOI: 10.12142/ZTETJ.202001014
- [6] GAO Z, DAI L L, WANG Z C, et al. Spatially common sparsity based adaptive channel estimation and feedback for FDD massive MIMO [J]. IEEE transactions on signal processing, 2015, 63(23): 6169 – 6183. DOI: 10.1109/TSP.2015.2463260
- [7] KUO P H, KUNG H T, TING P G. Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays [C]//Proceedings of 2012 IEEE Wireless Communications and Networking Conference. IEEE, 2012: 492 – 497. DOI: 10.1109/WCNC.2012.6214417
- [8] LU L, LI G Y, QIAO D L, et al. Sparsity-enhancing basis for compressive sensing based channel feedback in massive MIMO systems [C]//Proceedings of 2015 IEEE Global Communications Conference. IEEE, 2015: 1 – 6. DOI: 10.1109/GLOCOM.2015.7417036
- [9] DAUBECHIES I, DEFRISE M, DE MOL C. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint [J]. Communications on pure and applied mathematics, 2004, 57(11): 1413 – 1457. DOI: 10.1002/cpa.20042
- [10] KONG Q L, GONG R, LIU J T, et al. Investigation on reconstruction for frequency domain photoacoustic imaging via TVL3 regularization algorithm [J]. IEEE photonics journal, 2018, 10(5): 1 – 15. DOI: 10.1109/JPHOT.2018.2869815
- [11] METZLER C A, MALEKI A, BARANIUK R G. From denoising to compressed sensing [J]. IEEE transactions on information theory, 2016, 62(9): 5117 – 5144. DOI: 10.1109/TIT.2016.2556683
- [12] GAO Z, YAN S, ZHANG J, et al. ANN-based multi-channel QoT-prediction over a 563.4 km field-trial testbed [J]. Journal of lightwave technology, 2020, 38(9): 2646 – 2655
- [13] GAO Z G, ZHANG J W, YAN S Y, et al. Deep reinforcement learning for BBU placement and routing in C-RAN [C]//Proceedings of Optical Fiber Communication Conference (OFC). OSA, 2019: 1 – 3. DOI: 10.1364/ofc.2019.w2a.22
- [14] WEN C K, SHIH W T, JIN S. Deep learning for massive MIMO CSI feedback [J]. IEEE wireless communications letters, 2018, 7(5): 748 – 751. DOI: 10.1109/LWC.2018.2818160
- [15] WANG T Q, WEN C K, JIN S, et al. Deep learning-based CSI feedback approach for time-varying massive MIMO channels [J]. IEEE wireless communications letters, 2019, 8(2): 416 – 419. DOI: 10.1109/LWC.2018.2874264
- [16] LIU F, HE X C, LI C G, et al. CsiNet-plus model with truncation and noise on CSI feedback [J]. IEEE transactions on fundamentals of electronics, communications and computer sciences, 2020, E103.A(1): 376 – 381. DOI: 10.1587/transfun.2019eal2123
- [17] LU Z L, WANG J T, SONG J. Multi-resolution CSI feedback with deep learning in massive MIMO system [C]//Proceedings of ICC 2020 – 2020 IEEE International Conference on Communications. IEEE, 2020: 1 – 6. DOI: 10.1109/ICC40277.2020.9149229
- [18] LU Z L, ZHANG X D, HE H Y, et al. Binarized aggregated network with quantization: flexible deep learning deployment for CSI feedback in massive MIMO system [EB/OL]. [2021-10-01]. <https://ieeexplore.ieee.org/abstract/document/9684243>. DOI: 10.1109/TWC.2022.3141653
- [19] LIU L F, OESTGES C, POUTANEN J, et al. The COST 2100 MIMO channel model [J]. IEEE wireless communications, 2012, 19(6): 92 – 99. DOI: 10.1109/mwc.2012.6393523
- [20] STUBER G L, BARRY J R, MCLAUGHLIN S W, et al. Broadband MIMO-OFDM wireless communications [J]. Proceedings of the IEEE, 2004, 92(2): 271 – 294. DOI: 10.1109/JPROC.2003.821912
- [21] RASHEED M H, SALIH O M, SIDDEQ M M, et al. Image compression based on 2D discrete Fourier transform and matrix minimization algorithm [EB/OL]. [2021-10-01]. <https://www.sciencedirect.com/science/article/pii/S2590005620300096>
- [22] WANG Y S, YAO H X, ZHAO S C. Auto-encoder based dimensionality reduction [J]. Neurocomputing, 2016, 184: 232 – 242. DOI: 10.1016/j.neucom.2015.08.104
- [23] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84 – 90. DOI: 10.1145/3065386
- [24] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 770 – 778. DOI: 10.1109/CVPR.2016.90
- [25] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module [C]//Proceedings of the European Conference on Computer Vision. ECCV, 2018: 3 – 19. DOI: 10.1007/978-3-030-01234-2\_1
- [26] 3GPP. Study on channel model for frequencies from 0.5 to 100 GHz: TR 38.901 [S]. 2017

### Biographies

**GAO Zhengguang** (gao.zhengguang@zte.com.cn) received his BS degree from Hubei Engineering University, China in 2013, MS degree from South China Normal University, China in 2016, and PhD degree from Beijing University of Posts and Telecommunications, China in 2020. In his doctor's degree program, he was a visiting PhD student in High Performance Networks group, University of Bristol, UK from Nov. 1, 2018 to Nov. 1, 2019. After graduation, he was selected for "LAN JIAN" program of ZTE Corporation as an algorithm researcher. His current research interests include 5G/6G communication technologies, mobile networks, and machine learning for future communications.

**LI Lun** received his MS degree in electronics and communication engineering from Harbin Institute of Technology, China in 2018. He joined ZTE Corporation, China in 2018, where he is currently a technical pre-research engineer. His research interests include artificial intelligence/machine learning for wireless communications.

**WU Hao** received his BS degree from Beijing University of Posts and Telecommunications, China in 2010, and PhD degree from Southeast University, China in 2015, both in electrical engineering. He is now with the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, ZTE Corporation, where he is a senior expert on wireless communication research and standardization. During 2011 – 2012, he was a visiting student at Columbia University, USA. Since 2016, Dr. WU has been a delegate representing ZTE Corporation in 3GPP RAN and RAN1, to which he has submitted numerous contributions on 4G and 5G technologies including MIMO, UE power saving, positioning and so on. His research interests include MIMO wireless communications, antenna array systems, and signal processing.

**TU Xuezheng** is currently pursuing her master's degree with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China. She received her bachelor's degree in computer science and technology from Henan University, China in 2020. Her research interest is mainly communication-efficient distributed learning.

**HAN Bingtao** received his BS degree from Tianjin University, China in 2001, and MS degree from Nankai University, China in 2004. He is the deputy director of the State Key Laboratory of Mobile Network and Mobile Multimedia Technology, and the leader for "Adlik" project of the LF AI & Data Foundation. Currently, he is the AI system architect of Central R&D Institute, ZTE Corporation. His current research interests include deep learning algorithms, AI systems, and network intelligence. He is the author and co-author for numerous patents and related monographs.