Source Channel Model Selection for Downlink CSI Feedback Estimation using Transfer Learning in Massive MIMO System

Liu Jingyu*, Mondher Bouazizi†, Tomoaki Ohtsuki‡,
*Graduate School of Science and Technology, Keio Univerity, Yokohama, Japan
†Faculty of Science and Technology, Department of Information and Computer Science, Keio University, Yokohama, Japan
Email:*{liu, bouzizi}@ohtsuki.ics.keio.ac.jp,‡ohtsuki@ics.keio.ac.jp

***Abstract*—Transfer learning-based methods for Channel State Information (CSI) feedback can achieve a good CSI reconstruction performance with small amounts of computation time and cost. However, for higher efficacy, such methods need to fine-tune a properly selected source channel model each time. This is because fine-tuning a model trained on drastically different data will suffer from slow convergence and non-optimal performance, which defies the purpose of transfer-learning. In this paper, we use a one-stage algorithm utilizing k-nearest neighbors (k-NN) and Deep Neural Network (DNN) methods to classify the downlink Clustered Delay Line (CDL) channels into one of 5 channel models with the help of a selected set of features. Based on the classification results, the appropriate source channel model for fine-tuning is selected, and the model is fine-tuned on the received data. To demonstrate the usefulness of our method, we compare its results to those of existing methods. Through simulation, we show that the overall classification accuracy for 5 CDL channel models using the one-step algorithm is 80.73%. Upon classification, and by selecting the proper source model, our method can reach a better normalized mean square error (NMSE) of CSI feedback performance in a shorter time and with fewer samples.**

***Index Terms*—CDL, CSI feedback estimation, source model selection, deep transfer learning**

I. Introduction

Being one of the most important technologies of the fifth generation of mobile communication systems (5G), massive Multiple-Input Multiple-Output (MIMO) plays an important role in improving the transmission capacity and efficiency [1]. In a massive MIMO system, a large number of antennas are deployed at the Base Station (BS), allowing it to perform precoding and beamforming [2] [3]. These techniques, however, require acquiring an accurate knowledge of the Channel State Information (CSI) for them to be performed.

To obtain the CSI at the BS more accurately, much research based on the CSI feedback has been done [4]–[7]. However, the accuracy of conventional CSI feedback estimation methods exploiting the channel sparsity degrades with higher compression rates [4]. To address some of the issues of the conventional methods, a novel method called CsiNet [5] was introduced for the CSI feedback. In this method, deep learning with encoders and decoders was used to reconstruct the CSI at the BS. New improved neural networks such as CL-CsiNet and DFECsiNet iterated more on the ideas behind CsiNet, reaching better performance in reconstructing the CSI feedback [6] [7].

The problem with deep learning-based methods is that, when facing a new environment, the accuracy of the reconstructed CSI is not high. Therefore, new data samples from the new environment are required to retrain the deep learning model to achieve good performance. This action makes the training cost high. To solve this problem, transfer learning is introduced in CSI feedback [10]. Transfer learning is a popular deep learning-based technique that allows fine-tuning a source model trained on a certain channel model data with a few new samples [8]. In [10], transfer learning is used to reconstruct the CSI in the CDL channel model at BS. The Clustered Delay Line (CDL) channel model has two types of environment: Line-Of-Sight (LOS) and Non-Line-Of-Sight (NLOS). The issue with [10] is that only the CDL-A channel model is used as a source channel model for fine-tuning. Even when using transfer learning, to reach a good performance of reconstruction of the CSI, a long training time is needed. Based on [10], when unknown data are received from a given channel model, if a proper channel model is selected as a source to fine-tune with the received unknown channel data, better performance of reconstruction of the CSI can be achieved in shorter time [11] [12]. To achieve such results, a classification of the actual source channel is required for the received channel model data. Most of the existing works about channel classification so far were the LOS/NLOS environment classification and focused on the ultra-wideband (UWB) channels [9] [13].

The results of [11] [12] inspire this paper. They state that choosing a proper source channel model can reduce the training time and the number of training samples required. Our previous result [17] demonstrates that it is possible achieve a high accuracy for the LOS/NLOS classification. In this paper, we go a step further and use a set of selected features to do a multi-class classification by which we identify the CDL channel of the new environment. Using the results of the classification, we select the proper source channel model for fine-tuning the received data and perform the CSI feedback estimation through fine-tuning. In other words, the goal of this paper is to identify the CDL channel that fits the most with

This work was supported by JST ASPIRE Grant Number JPMJAP2326, Japan.

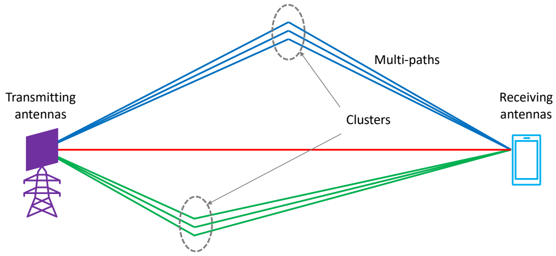

Fig. 1: The CDL channel model [14].

the new environment, and show that by doing such, better CSI feedback estimation can be achieved.

Concretely, we employ a one-stage algorithm utilizing k-nearest neighbors (k-NN) and Deep Neural Network (DNN) methods. Following the classification phase, we utilize the selected appropriate source channel model to perform the transfer learning method. This allows us to evaluate the performance of our approach in comparison to the methods used in previous research [10]. The results demonstrate that our method achieves better CSI feedback performance in a shorter time and while utilizing a reduced number of training samples.

II. Channel Model

We employ the CDL channel model, as specified in the 3GPP TR38.901 standard, for our study. The CDL channel models are designed to operate across a frequency range spanning from 0.5 GHz to 100 GHz, with a maximum bandwidth of 2 GHz. In our system, we consider a total of five distinct channel models: CDL-A, CDL-B, CDL-C, CDL-D, and CDL-E. The first three models correspond to various NLOS scenarios, while the latter two models LOS scenarios. The configuration of the CDL channel model is illustrated in Fig. 1. A few clusters are shown in the figure. Each cluster is a collection of paths with similar Angle Of Arrival (AOA) and Angle Of Departure (AOD). Each path can be NLOS or LOS as shown with the colors in the figure. In blue and green are two NLOS clusters where signals are reflected from some obstacles. In red is a path of LOS.

III. Proposed Method

A. Feature set

To perform the classification, we need to extract several features from the collected data. Therefore, we referred to our previous work [17] to identify which features can be used for this sake. We tested all the possible combinations of the candidate features. Based on the results of our early experiments [17], we selected a set containing four features, namely the kurtosis and the skewness of the sample data, the power [dB], and the maximum amplitude of each sample. These features exhibit the highest diversity between the different classes, making them high contributors to the classifications. While other features such as energy show a certain degree of diversity among the classes, they do not contribute as much to the classification, as they are one way or another derived from the selected ones. For instance, in Fig. 2, we show a

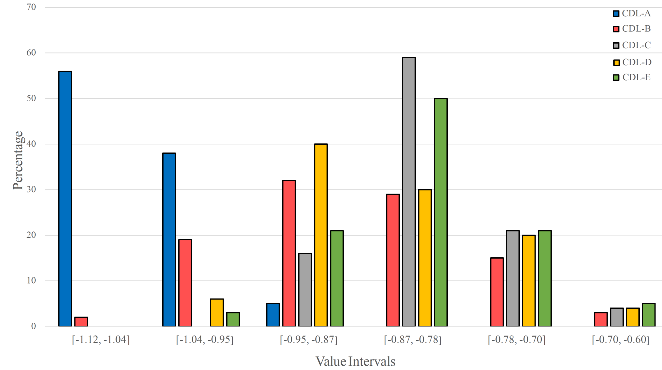

Fig. 2: Histogram of the distribution of kurtosis of sample data for each CDL channel.

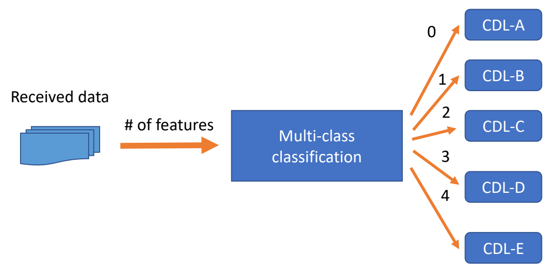

Fig. 3: The one-stage algorithm system for five CDL channels.

histogram of the distribution of the kurtosis for 100 randomly chosen samples in each CDL channel. The figure demonstrates that the kurtosis distribution for each channel of the CDL is distinct and varied. The kurtosis of the CDL-A channel is primarily distributed between -1.12 and -1.04, whereas the majority of samples for the CDL-B channel fall within the range of -1.04 to -0.78. For CDL-C, most samples range from -0.87 to -0.70. Between -0.95 and -0.70 is where the kurtosis of most CDL-D data samples is. For the last channel CDL-E, most data samples' kurtosis is distributed between -0.87 and -0.70 as over 70% of the points are distributed in this range. This information shows that the distribution of kurtosis of each channel is different. In other words, even though the samples from different CDL channels might overlap, the difference in the distribution is quite noticeable and can be used to justify the selection of this feature. The four features we finally use for our multi-class classification exhibit a similar pattern.

B. Classification Methods

We employ a one-stage algorithm for the classification, where we classify the received CDL channel data as one of the five channels directly. This process is summarized in Fig. 3. The classification is done using a k-NN classifier and a Deep Neural Network (DNN), as we will explain in detail later.

C. Numbers of samples

For the first round of tests, the number of samples is listed in Table I. Table I shows that for the multi-class classification 7500 samples of each CDL channel are used to train the classification model, and 2500 samples are used for the evaluation (i.e., testing). For the second test, we compare the

TABLE I: The number of samples for each simulation.

Simulation	Number of samples	
	Training the model	Testing
Multi-class classification	7500	2500
Comparison of two methods	500, 1000, 2500	2500
	4000, 5000	2500

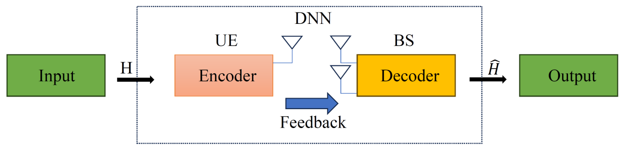

Fig. 4: Flow of the DNN-based downlink CSI feedback.

accuracy of the multi-class classification of the two algorithms with different numbers of training samples. As we can see we choose 500, 1000, 2500, 4000, and 5000 each time for each CDL channel.

D. Deep Neural Network for CSI Feedback

Fig. 4 illustrates the process of employing a DNN for downlink CSI feedback. The process commences with the reception of downlink CSI on the UE end. The received CSI matrix is denoted as $\mathbf{H}$. Subsequently, $\mathbf{H}$ is fed into the DNN's encoder, where it undergoes a transformation into a low-dimensional code word denoted as s.

$$s = f_{en}(\mathbf{H}), \tag{1}$$

where s is the code word and f_{en} is the encorder function. Next, s is sent back from the UE to the BS. After the BS receives s, s is input into the decoder of the DNN to reconstruct the original downlink CSI as

$$\hat{\mathbf{H}} = f_{de}(s) \tag{2}$$

where f_{de} is the decoder function. In this paper, the DNN is used as the backbone model to reconstruct the CSI feedback matrix.

E. Deep Transfer Learning

Transfer learning introduces two core concepts: domain and task. A domain is defined as $D = \{x, P(X)\}$, which includes the feature space x and the margin probability distribution $P(X)$, where $X = \{x_1, ..., x_n\} \in x$. The definition of the task is $T = \{y, f(\cdot)\}$, which includes the label space y and the target prediction function $f(\cdot)$, where $f(\cdot)$ can be learned from training data. Additionally, for the source model selection-based transfer learning method, we introduce the following 2 additional definitions [15]:

- Source data and model: Source data refers to the initial dataset used for the original training of the DNN. A source model, on the other hand, is a DNN model that has been trained using this source data.
- Target data and model: Target data is the dataset employed to refine the performance of the source model through fine-tuning. A target model is the resulting DNN that emerges from the fine-tuning process using the target data.

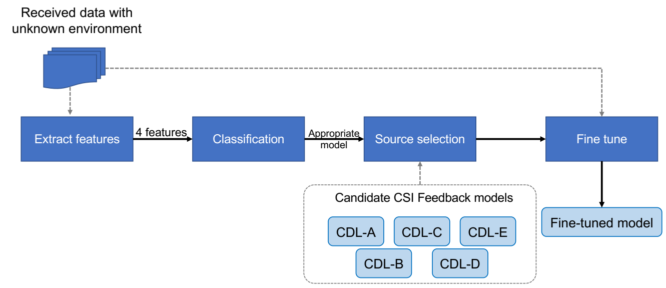

Fig. 5: Flowchart of the model selection and fine-tuning system.

IV. Source Selection Based Deep Transfer Learning

Prior research has utilized a substantial quantity of channel data samples to train a DNN from the beginning. Nevertheless, opting for this training approach for each distinct channel gives rise to elevated training expenses, encompassing both the consumption of data and time resources. Moreover, when facing a new environment or unknown environment, re-training the model each time will also cause high costs. Inspired by the evaluation from [12], we proposed this source model selection based transfer learning for downlink CSI feedback. The whole algorithm is shown in Fig. 5. This figure shows that at the beginning, the features are extracted from the data of the environment. We select the proper features to create a small feature set which we use for classification. This step is the most important step for our research. We need to select the source model based on our result of the classification. If the accuracy of the classification is high, we can select the proper source channel model for fine-tuning. If the accuracy of the classification is not high, the performance of CSI feedback will depend on the selected misclassified source model.

The NMSE is used to evaluate the performance of source model selection-based method, which is defined as follow:

$$\text{NMSE} = E[||\mathbf{H} - \hat{\mathbf{H}}||_2^2 / ||\hat{\mathbf{H}}||_2^2] \tag{3}$$

where $\mathbf{H}$ and $\hat{\mathbf{H}}$ represent the original and recovered CSI matrix, respectively.

V. Simulation Results and Evaluation

A. CDL channel

The parameters for our test are summarized in Table II. As shown in the table, during our simulations, we set the number of OFDM symbols N_o to 14. To analyze each CDL channel, 72 samples of (N_s) were required. The number of antennas at the BS (N_b) and that at the UE (N_u) are set to 32 and 2, respectively. The uplink and downlink frequencies are set to 2.0 GHz and 2.1 GHz, respectively. For each CDL channel, the Root Mean Square (RMS) delay spread is different. For CDL-A, the number is 129 ns, which is a normal delay. CDL-B and CDL-C have longer delay spread with an RMS equal to 634 ns. CDL-D and CDL-E are short delay channels, with

TABLE II: The basic simulation parameters of the CDL channels.

features	Values
Number of OFDM symbols (N_o)	14
Number of subcarriers (N_s)	72
Number of antennas at the BS (N_b)	32
Number of antennas at the UE side (N_u)	2
Uplink frequency	2.0 GHz
Downlink frequency	2.1 GHz
Velocities of the UE	4.8 km/h, 24 km/h 40 km/h, 60 km/h
Environment	Umi Street-canyon
RMS delay spread for CDL-A	129 ns
CDL-B and CDL-C	634 ns
CDL-D and CDL-E	65 ns

TABLE III: The accuracy of LOS/NLOS classification.

Method	**Accuracy**
Binary classification	96%
Two-step with DNN	96.53%
One-step with DNN	96.62%

their RMS delay spread set to 65 ns. Table II also shows that for our study, four different velocities are attributed to the UE. In each data sample, the UE is attributed a randomly selected velocity V_e from the four ones shown in the table.

B. Multi-class classification results and evaluation

To show the accuracy of our method, we also employ a two-stage classification algorithm for comparison. First, we use the two-step algorithm to classify each CDL channel. We proceed with a binary classification first, in which we aim to identify whether a given sample comes from a LOS or an NLOS environment. For this step, we use both k-NN and DNN methods for classification. Afterward, we use the one-step algorithm for multi-class classification for the five CDL channels. Moving forward, we calculated the LOS/NLOS accuracy of the classification results in Table III.

The results of multi-class classification for all the methods are presented in Table IV and Table V. From the two tables, we can see that the multi-class classification improves from the two-stage algorithm to the one-stage one. This means we can classify the CDL environment in a shorter time, while yielding better accuracy.

Next, we do the simulation of the two methods with different number of samples in order to know the least number of samples we need to have a stable result of classification. We opt for the one-stage algorithm as it presents results that are better than those of the two-stage algorithm while requiring less computation time and power. Fig. 6 and Fig. 7 shows the results of multi-class classification with different number of samples for each environment. Both figures show that the one-stage method yields a better accuracy of classification with different numbers of samples. From Fig. 6, the one-stage method can reach stable classification accuracy when more than 1500 samples are used for the NLOS environment. For LOS environmaent, the one-stage environment classification reaches a stable classification when more than 3000 samples are used as shown in Fig. 7. In conclusion, for the CDL channel multi-class classification, when using more than 3000 samples for each of the CDL channels, the classification is accurate and stable.

TABLE IV: The confusion matrix of the multi-class classification of the two-stage algorithm.

Actual label	**Classified as**				
	(A)	(B)	(C)	(D)	(E)
CDL-A	2137	32	71	201	59
CDL-B	75	1991	431	3	0
CDL-C	246	423	1811	2	18
CDL-D	91	1	6	2041	361
CDL-E	49	7	0	662	1782

TABLE V: The confusion matrix of the multi-class classification of the one-stage algorithm.

Actual label	**Classified as**				
	(A)	(B)	(C)	(D)	(E)
CDL-A	2226	31	78	123	42
CDL-B	54	2038	408	0	0
CDL-C	178	397	1899	4	22
CDL-D	114	1	6	2031	348
CDL-E	96	1	13	493	1897

C. CSI feedback performance based on source channel selection

After we select the source channel model, we fine-tune this model to the target CDL environment. The original (source) models are trained for 1000 epochs, whereas the fine-tuning is done for only 200 epochs. The reason for doing this is that all the source models are trained offline. When a model is selected for fine-tuning (after the classification), it needs to be fine-tuned in the shortest possible time. We fine-tune the target data with the different number of samples by using each CDL channel. Fig. 8 to Fig. 12 show the performance of CSI feedback of each CDL channel when fine–tuning with different numbers of samples. The number of samples we used for the experiments is summarized in Table VI. The proposed CSI feedback performance is represented by the dashed line. These five figures show that if the classification is correct, the ideal CSI feedback performance is the best.

Fig. 8 shows that when CDL-A is the target channel, besides CDL-A, CDL-B and CDL-C can give good performance and can even get close to the proposed performance when using 4000 samples for fine-tuning. The Table V shows that when mis-classification happens, the most possible mis-classified channel is CDL-D. From Fig. 8 when we use a CDL-D model to fine-tune a CDL-A environment (due to a misclassification), the performance is poor. However, Table V shows that for 3000 samples, only 123 samples (i.e., about 4.1% of the samples) will be mis-classified as CDL-D. Although using CDL-D to

TABLE VI: The number of samples for CSI feedback.

Simulation	**Number of samples**	
	Learning	Evaluation
Training the model	5000	2500
Fine-tuning	200, 500, 1000, 2000, 4000	2000

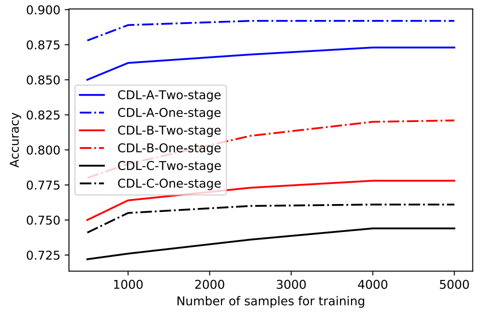

Fig. 6: The accuracy of multi-class classification of the two algorithms for NLOS CDL channels with different numbers of training samples.

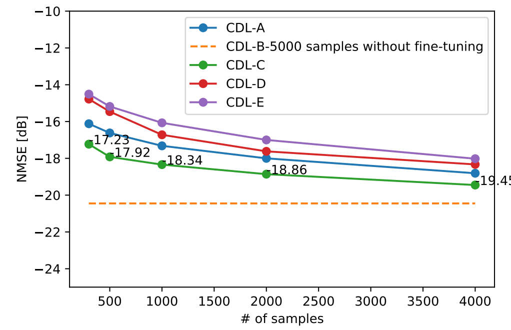

Fig. 9: CSI Feedback performance of CDL-B fine-tuned by each CDL channel data.

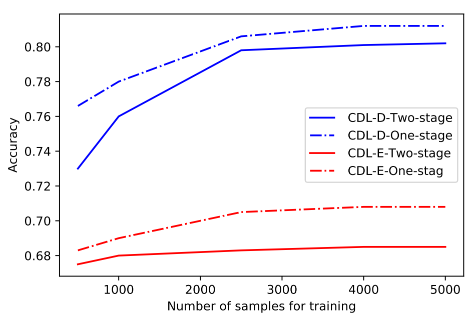

Fig. 7: The accuracy of multi-class classification of the two algorithms for LOS CDL channels with different numbers of training samples.

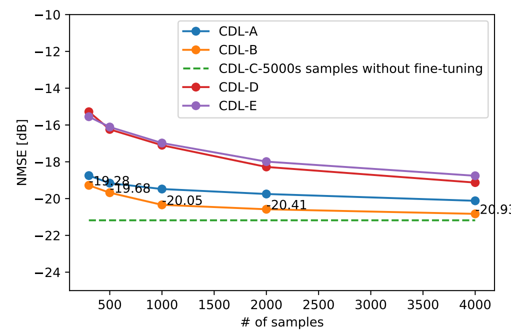

Fig. 10: CSI Feedback performance of CDL-C fine-tuned by each CDL channel data.

fine-tune CDL-A results in poor CSI feedback performance, the possibility is too low to happen.

Fig. 9 and Fig. 10 show the performance when CDL-B and CDL-C are the target environments. The confusion matrix given in Table V shows that CDL-B and CDL-C are mostly misclassified as one another, with 13.6% of CDL-B samples are misclassified as CDL-C and 13.2% the other way around. The possibility of misclassification of CDL-B and CDL-C is also not high. However, even when such misclassification happens, Fig. 9 and Fig. 10 show that the wrongly selected source model can be fine-tuned to reach a good CSI feedback performance.

Fig. 11 and Fig. 12 show the performance of CSI feedback when CDL-D and CDL-E are the target environment. Table V shows that when misclassification occurs, CDL-D will be mostly misclassified as CDL-E and CDL-E will be mostly misclassified as CDL-D with 11.6% and 15.4%, respectively. If the source model is wrongly selected due to misclassification, the performance is not very good. However, based on the majority voting rule, the accuracy of classification is high enough to select the correct source channel model.

These five figures show that CDL-B and CDL-C can have better CSI feedback performance no matter what the target

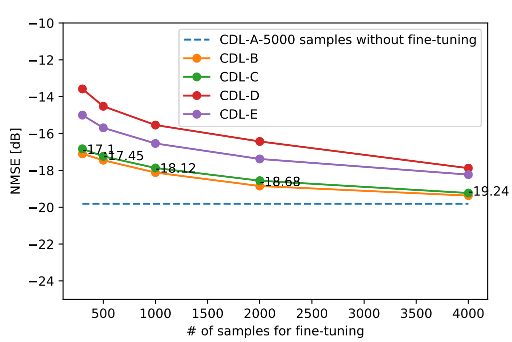

Fig. 8: CSI Feedback performance of CDL-A fine-tuned by each CDL channel data.

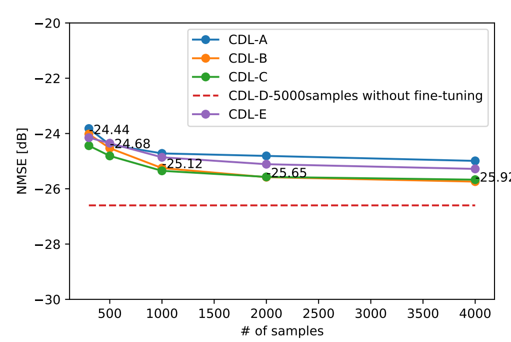

Fig. 11: CSI Feedback performance of CDL-D fine-tuned by each CDL channel data.

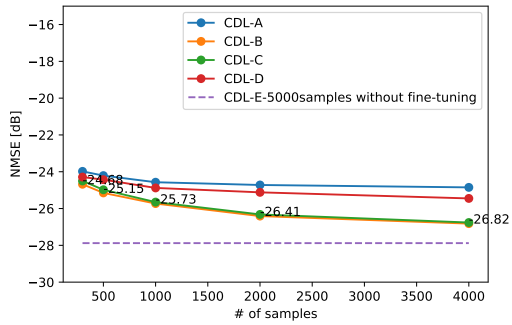

Fig. 12: CSI Feedback performance of CDL-E fine-tuned by each CDL channel data.

TABLE VII: Number of samples for each approach.

Approach	Number of samples		
	Training	Validation	Fine-tuning
CsiNet [5]	4000	1000	–
Transfer [8]	2000	500	1000
SM[1]selection	2000	500	1000

[1] Source model

environment is. However, the ideal source model still reaches the best CSI feedback performance.

D. Comparison

In this part, we compare the NMSE performances of CsiNet [5], transfer learning [8], and our proposed method. Table VII shows the number of samples we used for training the data and fine-tuning the target data.

Tables VII and VIII show that by using fewer samples, our method reaches the best CSI feedback performance in all scenarios. Compared with CsiNet [5] and the transfer learning method [8], our source selection-based transfer learning has a more obvious advantage under the same compression ratio with fewer samples. This is because our high classification accuracy lets us select the proper source channel for fine-tuning. This advantage allows our method to be implemented in real-world where only a few samples from an unknown environment are received. The results show that our method achieves better CSI feedback performance with fewer samples.

VI. Conclusion

In this paper, in order to identify the appropriate CDL source channel for transfer learning-based CSI feedback, we have proposed a novel method for classifying the actual CDL channel when given a sample of data collected from an unknown environment. The proposed method consists of a single-stage classification algorithm. In this one-stage algorithm, we classify the five channels directly. After the classification, we select the proper source model based on the classification result for fine-tuning the target data. The performance of CSI feedback of our method can reach a better performance than that of the previous research by using fewer samples.

TABLE VIII: Comparison of the CSI feedback performance.

Scenario	Number of samples CR[1]= 1/8		
	CsiNet (dB) [5]	Transfer (dB) [8]	SM[2]Selection (dB)
CDL-A	-14.38	–	-17.99
CDL-B	-11.35	-8.44	-18.43
CDL-C	-13.51	-13.69	-18.31
CDL-D	-16.38	-14.67	-20.41
CDL-E	-15.16	-14.13	-19.97

[1] Compression ratio
[2] Source model

References

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

[2] J. Kim, J. Choi, and J. M. Cioffi, “Cooperative Distributed Beamforming With Outdated CSI and Channel Estimation Errors,” in *IEEE Trans. Commun.*, vol. 62, no. 12, pp. 4269–4280, Dec. 2014.

[3] S. Qiu, Da Chen, D. Qu, K. Luo and T. Jiang, “Downlink Precoding With Mixed Statistical and Imperfect Instantaneous CSI for Massive MIMO Systems,” in *IEEE TVT* vol. 67, no. 4, pp. 3028–3041, Apr. 2018.

[4] P. Kuo, H. T. Kung and P. Ting, “Compressive sensing based channel feedback protocols for spatially-correlated massive antenna arrays,” in *Proc. IEEE WCNC*, pp. 492–497, apr. 2012.

[5] C. Wen, W. Shih and S. Jin, “Deep Learning for Massive MIMO CSI Feedback,” in *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 748–751, Oct. 2018.

[6] S. Ji and M. Li, “CLNet: Complex Input Lightweight Neural Network Designed for Massive MIMO CSI Feedback,” in *IEEE Wireless Commun. Lett.*, vol. 10, no. 10, pp. 2318–2322, Oct. 2021.

[7] Y. Xu, M. Zhao, S. Zhang and H. Jin, “DFECsiNet: Exploiting Diverse Channel Features for Massive MIMO CSI Feedback,” in *Proc. 13th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, pp. 1–5, Oct. 2021.

[8] Y. Yang, F. Gao, Z. Zhong, B. Ai and A. Alkhateeb, “Deep Transfer Learning-Based Downlink Channel Prediction for FDD Massive MIMO Systems,” in *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7485–7497, Dec. 2020.

[9] A. Kirmaz, D. S. Michalopoulos, I. Balan and W. Gerstacker, “LOS/NLOS Classification Using Scenario-Dependent Unsupervised Machine Learning,” *2021 IEEE PIMRC*, pp. 1134–1140, Sep. 2021.

[10] J. Zeng et al., “Downlink CSI Feedback Algorithm With Deep Transfer Learning for FDD Massive MIMO Systems,” in *IEEE Trans. Cogn. Commun. Netw.*, vol. 7, no. 4, pp. 1253–1265, Dec. 2021.

[11] M. Inoue and T. Ohtsuki, “Evaluation of Source Data Selection for DTL Based CSI Feedback Method in FDD Massive MIMO Systems,” *IEICE Tech. Rep.* RCS2022–62, Ryukyu Univ., Jun. 2022 (in Japanese).

[12] M. Inoue, T. Ohtsuki, K. Yamamoto, and G. Gui, ”Evaluation of Source Data Selection for DTL Based CSI Feedback Method in FDD Massive MIMO Systems,” *IEEE Consum. Commu. Netw. Conf. (CCNC2023)*, Las Vegas, NV, Jan. 2023.

[13] C. Jiang, J. Shen, S. Chen, Y. Chen, D. Liu and Y. Bo, “UWB NLOS/LOS Classification Using Deep Learning Method,” in *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2226–2230, Oct. 2020.

[14] 3GPP TD RP-151606: “Study on channel model for frequency spectrum above 6 GHz”.

[15] S. J. Pan and Q. Yang, ”A Survey on Transfer Learning,” in IEEE Trans. on Knowledge and Data Engineering, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.

[16] C. Tan, F.Sun, T.Kong, W.Zhang, C. Yang, and C. Liu: ”A survey on deep transfer learning,” in Proc. 27th Int. Conf. Artif. Neural Netw., pp. 270-279, Oct.2018.

[17] J. Liu, M. Bouazizi, and T. Ohtsuki, ”LOS/NLOS Classification for Downlink CDL Channel Using Supervised Learning.” IEICE Proceedings Series 72.O2-4 (2022).