C2R-KD: Complex to Real Knowledge Distillation (Student Abstract)

Byunghyuk Youn, Ohyun Jo*

Chungbuk National University, Department of Computer Science, 28644, Cheongju, Republic of Korea
{ybh0311, ohyunjo}@chungbuk.ac.kr

Abstract

In this work, C2R-KD is proposed, applying a Complex-to-Real projection to map complex domain features into the real domain. C2R-KD mitigates complex-real domain mismatch to strengthen the representational capacity of the student model and further improves the knowledge distillation model performance through the hybrid distillation of features and logits simultaneously. Experimental result demonstrates higher accuracy than the conventional KD across all test environments.

Introduction

Deep learning (DL) has attracted growing interest in signal processing due to its capability to extract representations directly from raw signals (Kim and Jo 2024). Nevertheless, real-world wireless signals are typically represented as complex-valued sequences, while most DL models are tailored for real-valued inputs. This mismatch often results in information loss and degraded performance when applied to complex signals (Kim et al. 2025). Meanwhile, complex-valued neural networks (CVNN) have been employed to mitigate this mismatch, offering a more natural treatment of phase and amplitude information in signals (Tu et al. 2020). However, CVNN typically has large number of parameters and high computational complexity, which limits their practical deployment(Xiao, Yang, and Feng 2023).

To overcome limitations, This study introduces a novel knowledge distillation(KD) framework. The method distills both logit- and feature-level knowledge from a teacher and a Teacher Assistant (TA). Stepwise distillation reduces the performance gap between large and lightweight models and improves training stability. By preserving features of complex-valued signals and the amplitude–phase structure, the proposed framework enhances signal processing performance. Furthermore, projecting complex features into the real domain enables the use of a lightweight student model based on real-valued operations, thereby significantly reducing the parameter count of the model.

[*]Corresponding author

C2R-KD

Complex-Valued Neural Network The CVNN processes complex data using complex-valued operation defined as:

$$f(\mathbf{z}) = (\mathbf{W^R z^R} - \mathbf{W^I z^I}) \; + \; j(\mathbf{W^I z^R} + \mathbf{W^R z^I}). \quad (1)$$

where W denotes the weight matrix and $\mathbf{z}$ the complex-valued signal. R and I indicate the real and imaginary parts. f denotes a complex-valued operation. CVNN preserves interactions between the real and imaginary components.

Complex to Real Feature Projection CVNN incurs elevated computational cost and larger model size due to complex arithmetic. To improve the efficiency–performance trade-off under resource constraints, KD has been proposed and increasingly adopted (Youn and Jo 2025). Complex–real domain mismatch, specifically disparities in numerical representations and operational semantics, preclude direct alignment of the representation spaces. As a result, feature-level knowledge distillation cannot be performed. To enable KD, complex feature maps are transformed into real-valued representations.

$$\Phi(\mathbf{z}) = \{\, |\mathbf{z}_n| \;\mid\; \mathbf{z}_n \in \mathbf{z} \,\}. \quad (2)$$

where $\mathbf{z}_n$ represents an individual complex component contained within the feature map. $\Phi(\mathbf{z})$ represents the amplitude of the complex feature map. The Complex to Real Feature Projection converts complex-valued features into the real domain while preserving amplitude information that reflects the signal's energy distribution and intensity. The complex-valued model can produce a stable and compact real-valued representation, enabling the student model to inherit spatial correlations and other structural properties embedded in the complex domain.

Consequently, this projection mitigates the representational discrepancy between complex and real domains and facilitates consistent and efficient knowledge transfer during distillation, thereby enhancing the stability of convergence and the generalization performance of the model. Furthermore, it enables the realization of Complex-to-Real Knowledge Distillation (C2R-KD), which effectively transfers the knowledge of a complex-valued teacher model to a real-valued student model.

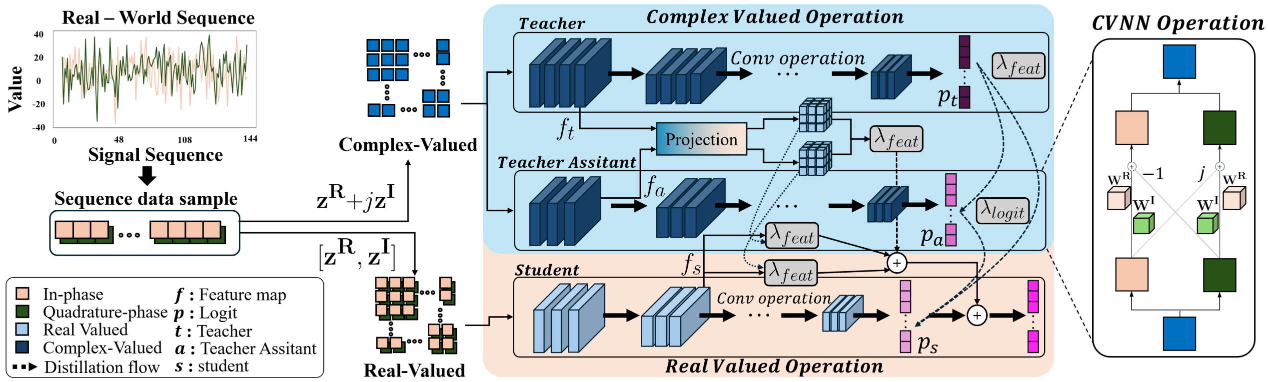

Figure 1: Framework of C2R-KD : Complex to Real Knowledge Distillation

Signal to Noise Ratio	Signal Classification Accuracy (%)				
	C2R-KD	Feature KD	Logit KD	CVCNN	RVCNN
-2.51 dB	**99.18**	98.31	96.18	98.69	98.30
-2.74 dB	**99.06**	98.61	95.59	98.74	98.08
-2.81 dB	98.98	98.19	**99.36**	98.69	97.94
-2.99 dB	**98.19**	96.73	97.44	97.54	97.38
-3.13 dB	**97.70**	96.38	95.15	96.98	95.93
-3.42 dB	**95.21**	92.85	92.34	93.69	92.98
-3.70 dB	**88.01**	84.83	83.84	86.60	84.23
-4.11 dB	**66.34**	59.18	45.19	65.49	55.85

Table 1: Performance Comparison of Distillation Methods

Hybird Loss function for C2R-KD C2R-KD employs a multi-level distillation scheme, where the teacher and TA models collaboratively transfer knowledge to the student model. C2R-KD achieves complementary guidance through both logit- and feature-level alignment. The teacher and TA models adopt CVNN architectures, reflecting higher expertise and capacity to handle greater complexity and computational demands, whereas the student model, designed using only a lightweight real-valued operations, leverages the distilled knowledge to achieve improved accuracy with significantly reduced computational cost.

As the C2R-KD framework in Figure 1 supports multiple distillation pathways, complementary hybrid distillation losses are designed to jointly capture features and logits. The overall training objective is given by:

$$\mathcal{L} = \sum_{k \in \mathcal{K}} \alpha_k \, \lambda_{\text{feat}}(T_k \mid S_k) + \sum_{k \in \mathcal{K}} \beta_k \, \lambda_{\text{logit}}(T_k \mid S_k) + \mathcal{L}_{\text{CE}}. \tag{3}$$

where $\mathcal{K}$ denotes the set of upper to lower distillation pairs (e.g., $Teacher \to TA$, $Teacher \to Student$, $TA \to Student$). For each $k \in \mathcal{K}$, T_k denotes the output of the upper model(e.g., $Teacher$,TA) and S_k denotes the output of the lower model(e.g., TA , $Student$). λ_{feat}, λ_{logit} are the feature and logit-level distillation losses. α_k and β_k are their respective weights. $\mathcal{L}_{\text{CE}}$ is the cross-entropy loss used to supervise the student. Hybrid distillation loss integrates feature and logit KD with cross-entropy supervision and improves training stability through complementary loss pathways.

Experiment

Experiment Setup In the experiment, DMRS (DeModulation Reference Signal) data collected in a real 5G environment was utilized (Han, Jo, and Kim 2022). The signals were collected at eight noise levels using an implemented 5G testbed, and the models were evaluated separately under each condition. The experiments employed VTCNN2 (O'shea and West 2016) as the Teacher, CVCNN (Wang et al. 2021) as the TA, and RVCNN—the real-valued version of CVCNN—as the Student.

To validate the effectiveness of C2R-KD, signal classification experiments employing Logit and Feature KD were conducted. Performance was compared against Logit KD, Feature KD, the Student without distillation, and the TA models.

Evaluations Table 1 shows the signal classification performance results. C2R-KD consistently exhibited better classification accuracy across compared to benchmarks. In particular, it showed an improvement of 10.49%p compared with the baseline student model RVCNN at lowest SNR of -4.11dB.

Additionally, C2R-KD achieved higher accuracy than the CVNN-based TA model. This demonstrates that C2R-KD effectively transfers the benefits of complex representations to the student in form of real-valued knowledege distilled form both teacher and TA. thereby improving performance, and further suggests that real-valued models can surpass complex-valued models.

Conclusion

In this study, C2R-KD is proposed as a cross-domain knowledge distillation that transfers knowledge from the complex domain to the real domain through domain transformation and by combining features and logits. The proposed approach was validated using real-world collected data and achieved higher accuracy than complex-valued models by employing a real-valued student model with C2R-KD. In future work, the framework will be extended by incorporating multiple TAs for knowledge transfer, while further improving overall performance.

Acknowledgments

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT). (No.2021R1A2C2095289) In part, this work was supported by the Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT). (No. 2021-0-00165, Development of 5G+ Intelligent Basestation Software Modem).

References

Han, J. Y.; Jo, O.; and Kim, J. 2022. Exploitation of channel-learning for enhancing 5G blind beam index detection. *IEEE Transactions on Vehicular Technology*, 71(3): 2925–2938.

Kim, J.; and Jo, O. 2024. MuShAug: Boosting Sequence Signal Classification via Multishape Augmentation. *IEEE Internet of Things Journal*, 11(20): 32585–32597.

Kim, J.; Min, W.; Kim, J.; and Jo, O. 2025. ComplexRep : Integrating Learned Representations to Enhance Complex-valued Data Transparency. *IEEE Internet of Things Journal*, 1–1.

O'shea, T. J.; and West, N. 2016. Radio machine learning dataset generation with gnu radio. In *Proceedings of the GNU radio conference*, volume 1.

Tu, Y.; Lin, Y.; Hou, C.; and Mao, S. 2020. Complex-valued networks for automatic modulation classification. *IEEE Transactions on Vehicular Technology*, 69(9): 10085–10089.

Wang, Y.; Gui, G.; Gacanin, H.; Ohtsuki, T.; Dobre, O. A.; and Poor, H. V. 2021. An efficient specific emitter identification method based on complex-valued neural networks and network compression. *IEEE Journal on Selected Areas in Communications*, 39(8): 2305–2317.

Xiao, C.; Yang, S.; and Feng, Z. 2023. Complex-valued depthwise separable convolutional neural network for automatic modulation classification. *IEEE Transactions on Instrumentation and Measurement*, 72: 1–10.

Youn, B.; and Jo, O. 2025. Transferring Complex-Domain Representations to Real-Valued Networks through Knowledge Distillation. In *26th International Symposium on Advanced Intelligent Systems*, volume 1.