Efficient Automatic Modulation Classification for Resource-Limited Devices

Pedro Marcio Raposo Pereira

April 15, 2024

1 Research Proposal

Military applications for Automatic Modulation Classifier (AMCs) are crucial in the realm of Electronic Warfare (EW) as they enhance situational awareness and enable prompt responses to hostile transmissions. AMCs acts as an intermediary process between signal detection and data demodulation or system reaction within EW operations [1]. Its significance lies in its ability to identify and classify signals without prior knowledge of specific signal parameters such as carrier frequency, phase offsets, or bandwidth. One primary military application of AMC lies in its role within EW equipment. Military forces rely on effective signal classification to detect, locate, and identify enemy transmissions swiftly and accurately. This capability is essential for generating timely and precise jamming responses or deploying countermeasures against hostile signals. In such scenarios, the efficiency and speed of the classification algorithm are critical factors, as any latency could compromise the effectiveness of EW operations.

AMC systems typically consist of two main components: signal preprocessing and the classification algorithm. The signal preprocessing stage involves estimating various signal parameters such as power, signal-to-noise ratio, time of arrival, pulse width, or carrier frequency. The accuracy of these estimations directly impacts the performance of the classification algorithm. Therefore, a joint design approach that optimizes both signal preprocessing and the classification algorithm is preferred to ensure optimal performance.

In the realm of classification algorithms, there are two primary approaches: likelihood-based (LB) and feature-based (FB). LB algorithms, also known as decision-theoretic approaches, leverage the likelihood function of received signals to optimize classification. While theoretically optimal, LB algorithms often entail high computational complexity, leading to practical challenges in real-time implementation [2–4]. Consequently, simplified or sub-optimal versions of LB algorithms are commonly used in military applications to balance computational demands with classification accuracy.

In contrast, FB algorithms adopt a pattern recognition approach by extracting signal features and processing them in a predefined manner for classification. Although not theoretically optimal, FB algorithms offer simpler implementation and lower computational overhead [3], making them suitable for real-time applications in military environments where speed and efficiency are paramount.

Deep Learning (DL), particularly in FB approaches within AMC, has garnered attention for its ability to achieve robust classification performance while maintaining low computational costs compared to traditional LB approaches. Firstly, a novel data preprocessing method aimed at enhancing the efficiency of convolutional neural networks (CNN)-based classification by optimizing the form of signal samples is proposed in [5], resulting in a significant 10% improvement in accuracy compared to traditional CNN approaches. Additionally, a CNN architecture incorporating residual blocks is proposed, achieving a maximum accuracy of 93.7% on the RadioML2016.10a dataset under challenging signal-to-noise ratio conditions, surpassing the performance of existing state-of-the-art classifiers in automatic modulation classification.

In the context of resource-constrained devices, a high-efficiency AMC is proposed in [6]. Leveraging stacking quasirecurrent neural network (S-QRNN) layers for feature extraction, the architecture integrates convolutional layers for low-latency feature extraction and a minimalist recurrent pooling function for enhanced classification accuracy. Evaluation against state-of-the-art classifiers demonstrates superior efficiency, with the proposed S-QRNN classifier exhibiting, on average, a 75.83% higher efficiency and a 26.54% lower execution latency. Additionally, by introducing gated recurrent units (GRUs) for temporal feature extraction, the resulting GS-QRNN classifier achieves an average efficiency increase of 191% compared to existing models, with a 59.31% lower execution latency on average. These findings highlight the effectiveness of the proposed architectures in meeting CR-IoT device requirements while maintaining high classification accuracy and efficiency.

Signal-based convolutional neural networks (SBCNN) [7] and Image-based convolutional neural networks (IBCNN) [8] represent two distinct approaches within the realm of DL-based AMC. The integration of these models in a hybrid DL framework offers a comprehensive solution for enhancing classification accuracy and robustness.

In SBCNN, the focus lies on analyzing the performance of CNN with varying convolution filter sizes to achieve optimal classification results. By exploring different filter sizes, SBCNN aims to extract relevant features from preprocessed signal data, enhancing the discriminative power of the classifier. The extracted features from pre-training SBCNN are then transformed into images to train the IBCNN model. This integration of SBCNN and IBCNN capitalizes on the strengths of both signal-based and image-based approaches, leveraging the complementary nature of signal and image representations in modulation classification tasks.

While the integration of SBCNN and IBCNN in a hybrid deep learning framework offers potential benefits in terms of classification accuracy, it also introduces certain challenges, particularly concerning model size and processing capability.

One notable concern with the adoption of a double-stage CNN approach is the increase in model size. Combining SBCNN and IBCNN into a single framework results in a larger and more complex model architecture. The increased number of parameters and layers may demand higher processing capability, both in terms of computational power and memory resources. In scenarios where computational resources are limited, deploying such large models could pose significant challenges, potentially leading to performance degradation or operational constraints.

Moreover, the larger model size can exacerbate system latency, which is a critical factor in many real-time applications, including wireless communication systems. The additional computational burden imposed by the double-stage CNN may lead to longer processing times, delaying the classification of incoming signals. In time-sensitive environments such as military operations, even minor increases in latency can have detrimental effects on system performance and responsiveness.

To address the challenges posed by the deployment of AMC in limited computational scenarios, the utilization of response-based knowledge distillation (KD) emerges as a promising solution. Response-based KD offers a methodology to train compact yet accurate AMC models by transferring knowledge from larger, more complex models to smaller ones. By optimizing the student model through a loss function that evaluates the disparity between its outputs and those of the teacher model, response-based KD enables the creation of smaller models while preserving crucial aspects of accuracy. The distillation loss, a key component of response-based KD, is computed based on the dissimilarity between the probability distributions of the teacher's and student's predictions. By applying temperature scaling to soften the logits and balancing the contributions of both hard and soft targets through appropriate weighting, the distillation loss facilitates effective knowledge transfer [9, 10]. This approach allows for the creation of smaller AMC models that are better suited for deployment on resource-constrained devices, without sacrificing significant accuracy. Additionally, response-based KD can be further optimized to improve latency, ensuring efficient inference in AMC tasks within edge computing scenarios.

References

[1] V. Iglesias, J. Grajal, and O. Yeste-Ojeda, "Automatic modulation classifier for military applications," in *2011 19th European Signal Processing Conference*, 2011, pp. 1814–1818.

[2] Y. Yuan, P. Zhao, B. Wang, and B. Wu, "Hybrid maximum likelihood modulation classification for continuous phase modulations," *IEEE Communications Letters*, vol. 20, no. 3, pp. 450–453, 2016.

[3] X. Yan, G. Zhang, and H.-C. Wu, "A novel automatic modulation classifier using graph-based constellation analysis for M -ary qam," *IEEE Communications Letters*, vol. 23, no. 2, pp. 298–301, 2019.

[4] A. Kumar, S. Majhi, G. Gui, H.-C. Wu, and C. Yuen, "A survey of blind modulation classification techniques for ofdm signals," *Sensors*, vol. 22, no. 3, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/1020

[5] H. Zhang, M. Huang, J. Yang, and W. Sun, "A data preprocessing method for automatic modulation classification based on cnn," *IEEE Communications Letters*, vol. 25, no. 4, pp. 1206–1210, 2021.

[6] P. Ghasemzadeh, M. Hempel, and H. Sharif, "Gs-qrnn: A high-efficiency automatic modulation classifier for cognitive radio iot," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 9467–9477, 2022.

[7] S.-H. Kim, J.-W. Kim, V.-S. Doan, and D.-S. Kim, "Lightweight deep learning model for automatic modulation classification in cognitive radio networks," *IEEE Access*, vol. 8, pp. 197 532–197 541, 2020.

[8] S.-H. Kim, C.-B. Moon, J.-W. Kim, and D.-S. Kim, "A hybrid deep learning model for automatic modulation classification," *IEEE Wireless Communications Letters*, vol. 11, no. 2, pp. 313–317, 2022.

[9] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," 2015.

[10] Z. Li, H. Li, and L. Meng, "Model compression for deep neural networks: A survey," *Computers*, vol. 12, no. 3, 2023. [Online]. Available: https://www.mdpi.com/2073-431X/12/3/60