Deep Learning-Based Modulation Recognition with Multi-Scale Temporal Feature Extraction

Da-Min Shin
Dept. of Electrical and Electronics Engineering
Pusan National University
Busan, Republic of Korea
ekals2020@pusan.ac.kr

Do-Hyun Park
Dept. of Electrical and Electronics Engineering
Pusan National University
Busan, Republic of Korea
dohpark@pusan.ac.kr

Hyoung-Nam Kim
School of Electrical and Electronics Engineering
Pusan National University
Busan, Republic of Korea
hnkim@pusan.ac.kr

***Abstract*—This paper studies a deep learning-driven method for identifying modulation types in communication signals without prior information. To effectively capture the relationship between points in the time-series data, we apply a multi-head self-attention mechanism. The results demonstrate superior modulation recognition accuracy compared to conventional AMR models. In particular, the classification accuracy between QAM16 and QAM64 was significantly improved, confirming that the performance of the model has been enhanced compared to previous models.**

***Index Terms*—automatic modulation recognition, multi-head self-attention**

I. Introduction

Automatic modulation recognition (AMR) is a cornerstone of electronic warfare support, where high accuracy is crucial for successful signal classification and operational efficiency. It intercepts unknown communication signals and estimates the modulation scheme [1]. As unmanned aerial vehicles (UAVs) play a key role in contemporary warfare [2], the importance of AMR in estimating hostile aircraft information has grown significantly.

Traditional modulation recognition methods are categorized into likelihood-based techniques and feature-based techniques [3]. Likelihood-based approaches are theoretically optimal, but they demand substantial computational resources, making them impractical in some scenarios. Feature-based methods have the drawback of relying on manual extraction of features [4]. Recently, research on AMR has increasingly focused on methods driven by deep learning. Deep learning-based techniques eliminate the need for manual feature extraction and achieve superior accuracy in modulation recognition compared to feature-based methods.

In this paper, multi-head self-attention is applied for modulation recognition, enabling the deep learning-based AMR model to capture the relationship between features. To analyze the significance of the attention mechanism in AMR, we compare the classification accuracy of the algorithms across different signal-to-noise ratio (SNR) conditions. A confusion matrix is used to analyze the error types for each modulation scheme.

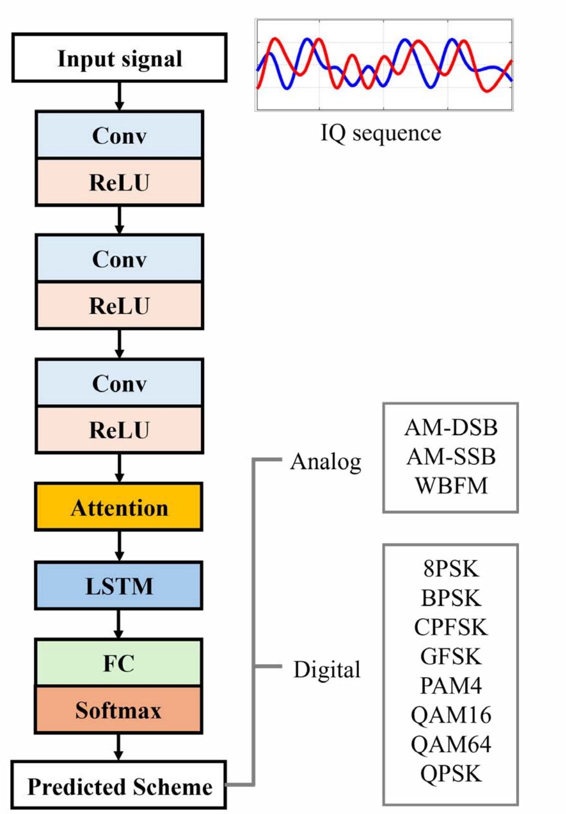

Fig. 1. Structure of the attention-adapted network.

II. Multi-Head Self-Attention Mechanism

The multi-head (MH) attention mechanism [5] is a technique that focuses on the important features of given data by utilizing the correlations between input data. The mathematical representation of the attention mechanism is given below:

$$\text{Attention}(Q, K, V) = \text{Softmax}(\frac{QK^T}{\text{dim}K})V \quad (1)$$

where the weight matrix is calculated using Q (query), K (key), and the dimensionality of the K. The output feature is then obtained by multiplying the weight matrix with V (value).

TABLE I
DESCRIPTION OF THE NEURAL NETWORK

Layer	Output size	Parameters
Conv1	[1024, 50, 120]	Kerner size = 9, ReLU
Conv2	[1024, 50, 112]	Kerner size = 9, ReLU
Conv3	[1024, 50, 104]	Kerner size = 9, ReLU
Attention	[1024, 104, 50]	# of heads h = 2
LSTM	[1024, 104, 250]	-
FC	[1024, 250]	ReLU
FC	[1024, 11]	Softmax

Notably, attention mechanism uses multiple heads to extract various features. Multi-head self-attention produces an output described as

$$\begin{aligned} \mathrm{MH}(Q,K,V) &= \mathrm{Concat}(\mathrm{head}_1,\ldots,\mathrm{head}_h)W^O, \\ \mathrm{head}_i &= \mathrm{Attention}(XW_i^Q, XW_i^K, XW_i^V). \end{aligned} \tag{2}$$

Each head consists of an attention block and generates three attention components from the input data X using different weight matrices W_i^Q, W_i^K, and W_i^V. After concatenating the attention maps from each head, the process employs a linear transformation with W^O to compute the final output feature. This multi-head self-attention mechanism supports the identification of latent features, allowing the model to analyze data from multiple perspectives.

III. Neural Network Architecture for AMR

Fig. 1 visualizes the structure of the neural network proposed in this work for modulation recognition. We adapted a model [6] designed for communication signal modulation recognition. To extract characteristics of input data, the architecture employs convolutional neural networks (CNN) [7] and subsequently utilizes long short-term memory (LSTM) [8]. In this paper, multi-head self-attention was applied to capture dependencies across time and channels for modulation recognition. Since communication signals exhibit patterns along the time axis, using multi-head attention to analyze correlations across different time ranges is crucial. Considering the length of the received signal and the computational load, two heads were used in each attention layer. The input size and parameters of each layer are shown in Table 1.

IV. Simulation Results

A. Dataset

We used the RadioML2016.10a dataset [9], which consists of 11 modulation schemes, including eight digital modulation techniques and three analog modulation techniques. The dataset contains 220,000 samples spanning SNR levels between –20 dB and 18 dB, with 2 dB increments. Each sample comprises an in-phase sequence and a quadrature sequence, both with a length of 128. The signals used for training were normalized with zero mean and unit variance.

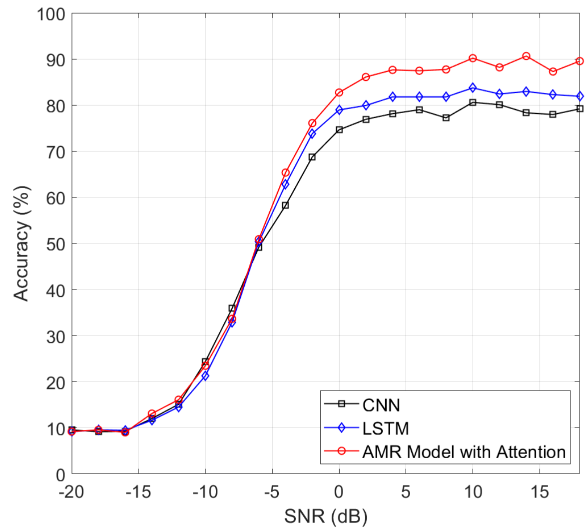

Fig. 2. Overall recognition performance comparison of the three models.

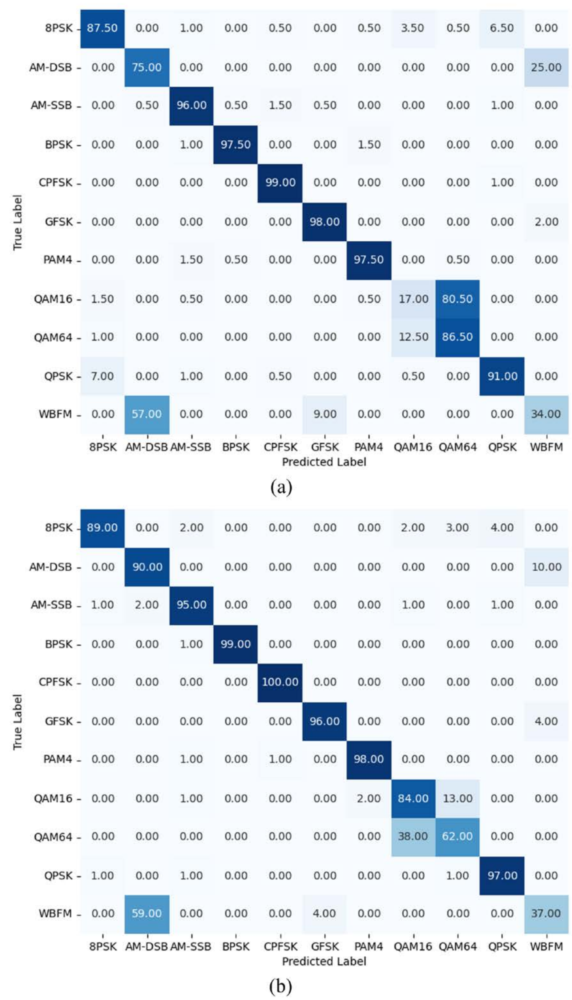

Fig. 3. Confusion matrices of models tested at 2 dB SNR (a) LSTM (b) attention-adapted model.

TABLE II
RECOGNITION ACCURACY ON RADIOML2016.10A (%)

SNR	CNN	LSTM	Attention-Adapted
0 dB	74.64	78.95	82.73
14 dB	78.32	82.95	90.64

B. Training hyperparameter

The training process employed the Adam optimizer [10] with a mini-batch size of 1024, epochs of 200, and an initial learning rate of 10^{-3}. To prevent overfitting, the model's learning rate was reduced by 20% every 10 epochs, and early stopping was applied if the validation loss did not decrease for 30 epochs.

C. Simulation results

Fig. 2 demonstrates the classification performance of each neural network with respect to SNR. The conventional AMR networks for comparison include CNN [6] and LSTM [6] networks. The attention-adapted model demonstrates superior performance in most SNR ranges compared to the other networks. Table 2 presents the modulation recognition performance at two SNR levels, where the attention-adapted model achieved the highest recognition rates of 82.73% at 0 dB and 90.64% at 14 dB.

Fig. 3 provides a view of the prediction results for the actual modulation schemes at 2 dB SNR. As shown in Fig. 3(a), the LSTM classifies many QAM16 signals as QAM64, with an average QAM classification accuracy of 51.75%. However, the application of the attention mechanism reduces the confusion between the QAM signals, achieving an average QAM classification accuracy of 73.00%, as shown in Fig. 3(b). These results indicate the improved robustness of the developed model in identifying QAM signals compared to conventional networks.

V. CONCLUSION

This paper analyzes a neural network that utilizes the multi-head self-attention mechanism to recognize communication signal modulation. This approach reduces the inductive bias caused by convolution, enabling the network to capture relationships across multiple time scales. Simulation results showed that the attention-adapted network outperformed other networks, achieving an accuracy of up to 90.64% and decreasing the confusion probability between specific modulation types. Our future efforts will focus on improving the classification accuracy of QAM signals and identifying more efficient ways to utilize the attention mechanism.

REFERENCES

[1] Y. Peng, L. Guo, J. Yan, M. Tao, X. Fu, Y. Lin, and G. Gui. "Automatic Modulation Classification Using Deep Residual Neural Network with Masked Modeling for Wireless Communications," *Drones*. vol.7, no. 6, pp. 390, 2023.

[2] U. Seidaliyeva, L. Ilipbayeva, K. Taissariyeva, N. Smailov, and E. T. Matson, "Advances and Challenges in Drone Detection and Classification Techniques: A State-of-the-Art Review," *Sensors*, vol. 24, no. 1, pp. 125, 2024.

[3] W. Xiao, L. Zhongqiang, and H. Qian. "A review of research on signal modulation recognition based on deep learning," *Electronics*, vol. 11, no. 17, pp. 2764, Sep. 2022.

[4] F. Meng, P. Chen, L. Wu and X. Wang. "Automatic Modulation Classification:A Deep Learning Enabled Approach," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 11, pp. 10760-10772, Nov. 2018.

[5] A. Vaswani et al., "Attention is all you need," *arXiv*, 2017.

[6] M. Zhang, Y. Zeng, Z. Han and Y. Gong, "Automatic Modulation Recognition Using Deep Learning Architectures," *International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 1-5, 2018.

[7] Y. LeCun et al, "Backpropagation applied to handwritten zip code recognition," *Neural Computation*, vol. 1, no. 4, pp. 541-551, Dec. 1989.

[8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997.

[9] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," *Proceedings of the GNU Radio Conference*, vol. 1, pp.1-6, 2016.

[10] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv*, 2014.