Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/121/2025.19842

Overview of Automatic Modulation Recognition Methods Based on Deep Learning

Shiqi Zhang[1,a,*]

[1]*School of Information Science and Technology, Beijing University of Technology, Beijing, 100124, China*
a. 18513819372@163.com
**corresponding author*

Abstract: Spectral resources are becoming increasingly scarce due to the rapid growth of communication data volumes, and the modulation methods of signals will also become more diverse. There is an urgent need for more effective modulation recognition methods. Firstly, based on a large amount of high-level information, this paper introduces shallow machine learning and some of its representative methods. Secondly, in response to the low accuracy and high computational complexity of most shallow machine learning models, this paper introduces the application of deep learning in the field of automatic modulation recognition. Finally, possible solutions to the existing problems and challenges in modulation recognition technology are proposed, and future prospects are discussed. Automatic Modulation Recognition(AMR) technology is relatively well developed in closed-set research, but it still needs to be strengthened in open-set research. In addition, research on unlabeled samples is also lacking, and there should be an increased focus on research for powerful hardware platforms capable of modulation recognition.

Keywords: Modulation Recognition, Deep Learning, AMR, Communication Signals.

1. Introduction

In traditional communication systems, the communicating parties are often in a cooperative relationship, familiar with each other’s modulation methods, which are usually singular. However, as communication technology continues to evolve, the modulation of information becomes increasingly complex and widespread in its applications. For example, in civilian fields, it is used for spectrum resource monitoring and optimization; in military fields, it is used for electronic countermeasures and communication reconnaissance during wartime[1-2]. When the communicating parties are not in a cooperative relationship but still wish to obtain information from each other, modulation recognition technology plays a significant role. Modulation recognition technology is a key technique situated between modulation and demodulation, with the core being pattern recognition, that is, categorizing signals based on the characteristics of the modulated signal waveforms. A basic pattern recognition system typically consists of four components: information acquisition, preprocessing, feature extraction, and classifier. Signals with different modulation methods can be identified and classified into specific modulation types after passing through the pattern recognition system. At the inception of modulation recognition, signal identification mainly relied on manual efforts, which were time-consuming and had low recognition efficiency. In 1969, C.S. Weaver and colleagues first proposed

DOI: 10.54254/2755-2721/121/2025.19842

Automatic Modulation Recognition (AMR) in a technical report at Stanford University, bringing the technology into public view. AMR refers to the technique of actively determining the modulation method of a signal from the received signal without prior knowledge of the modulation method used.

Modulation recognition technology can be divided into non-cooperative modulation recognition and cooperative modulation recognition based on the application scenario[3]. Compared to non-cooperative modulation recognition, cooperative modulation recognition requires more prior knowledge, such as a set of modulation methods, carrier frequency, pilot sequences, etc. Presently, AMR may be categorized into deep learning-based modulation recognition technology and non-deep learning-based modulation recognition technology. The primary techniques encompass likelihood ratio recognition methods grounded in decision theory and pattern recognition methods reliant on feature extraction. This article classifies basic AMR technologies based on whether they utilize deep learning, highlighting the characteristics of different AMR techniques, and points out some of the existing flaws in AMR technologies, such as the fact that most algorithms can only perform closed-set recognition for known modulation types and there is a lack of research data for unlabeled samples. The article also proposes some potentially viable research directions for existing problems, such as strengthening the study of open-set recognition.

2. Methods not based on deep learning

2.1. Likelihood ratio recognition method based on decision theory

The likelihood ratio recognition method refers to the transformation of the signal modulation recognition problem into a hypothesis testing problem based on probability theory and related theories of hypothesis testing. Initially, the maximum likelihood function is formulated based on the statistical properties of the signal model, from which the optimal decision threshold for the signal is derived. Subsequently, the likelihood ratio of the signal to be identified is compared with the threshold value, thus ascertaining the signal modulation type. Figure 1 illustrates the fundamental principle of modulation recognition utilizing the maximum likelihood function[4].

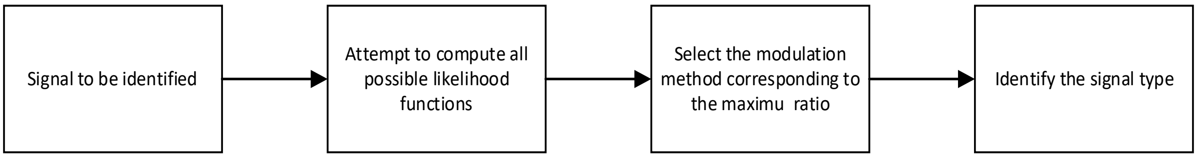


Figure 1: Flowchart of modulation recognition method based on likelihood function

This research method is mainly divided into: Average Likelihood Ratio Test (ALRT), Generalized Likelihood Ratio Test (GLRT), and Hybrid Likelihood Ratio Test (HLRT). The ALRT is a statistical detection method that assesses the validity of a hypothesis based on the likelihood ratio. This method is particularly applicable to the detection of burst signals in the fields of signal processing and communications[5]. The GLRT is a more universal likelihood ratio test method that allows parameters to be unknown under the alternative hypothesis and estimates these parameters through data. The Mixed Likelihood Ratio Test is a method that combines the characteristics of ALRT and GLRT. It can provide better performance in certain situations, especially when some signal characteristics are known and some are unknown[6].

However, this category of algorithms requires a lot of prior knowledge information, such as the mean, variance, covariance, and other prior information of the signal, which involves a large amount of calculation and poor universality. In non-cooperative communication, it is necessary to calculate the probability density function for each obtainable parameter, and the difficulty of obtaining

DOI: 10.54254/2755-2721/121/2025.19842

parameters is relatively high, having a low efficiency. Therefore, the number of scholars studying this method is gradually decreasing[7-9].

2.2. Pattern recognition method based on feature extraction

This method exploits the differences in spectral characteristics between different modulation schemes to extract various features of the signal. It then constructs a classifier to categorize this feature information, achieving the goal of accurately identifying unknown modulation methods. The principle is shown in Figure 2.

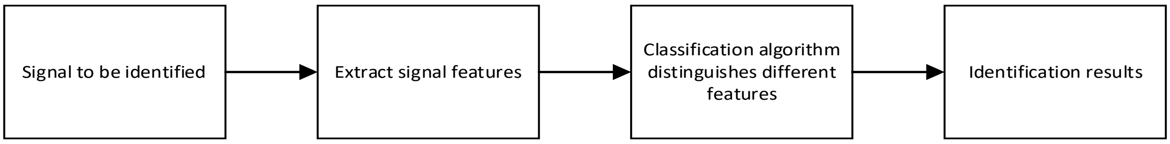


Figure 2: Flowchart of the recognition method based on feature extraction

Currently, the features used in feature extraction recognition methods mainly include:

- 1. Temporal domain characteristics including instantaneous amplitude, phase, and frequency;
- 2. Spectral domain characteristics such as cyclic spectrum, power spectrum, and higher-order cumulants[10];
- 3. Transform domain characteristics, including constellation diagrams and wavelet transforms[11].

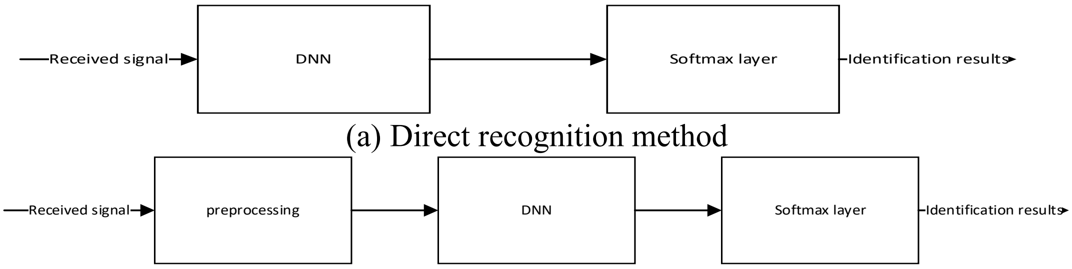
This algorithm has the advantages of being easy to simulate and implement, requiring less prior knowledge, having a smaller computational load compared to likelihood ratio functions, and being more efficient. Nonetheless, it imposes stringent demands on the retrieved features and requires the formulation of suitable feature parameters and classifiers with robust generalization abilities. Moreover, its recognition efficacy is constrained in situations with low signal-to-noise ratios.

3. Modulation recognition based on deep learning

Deep learning, as a rapidly developing field in machine learning in recent years, utilizes an end-to-end pattern and has shown unique technical advantages in many fields such as speech recognition, image recognition, natural language processing, and radar radiation source recognition, and has been widely applied. For example, the application of deep learning in language and speech recognition has a significant and far-reaching impact. Deep learning technology has significantly enhanced the efficacy of ASR systems, particularly in managing noise and various accents and dialects. Deep learning models may autonomously extract significant features from raw voice signals, obviating the necessity for manual feature engineering[12]. In addition, deep learning has also had a significant impact on image recognition. Deep learning models, such as AlexNet, have demonstrated higher accuracy in image classification tasks compared to traditional machine learning algorithms. Image classification is an application of machine learning technology, the goal of which is to train a model that can predict the category an input image belongs to. Deep learning models automatically learn high-level abstract information from samples through multiple hidden layers, which has a significant advantage over manual feature processing[13].

Modulation recognition methods based on deep learning can be divided into direct recognition methods and indirect recognition methods, as shown in Figure 3(a) and Figure 3(b).

DOI: 10.54254/2755-2721/121/2025.19842



(a) Direct recognition method

(b) Indirect recognition method

Figure 3: Deep learning-based recognition methods

The direct recognition method involves processing the RF signal received by the receiver, including down-conversion, and utilizing the time-domain complex baseband signal as the input for the neural network, training the network to perform AMR. The indirect recognition method entails preprocessing the time-domain complex baseband signal to get alternative transformed representations prior to inputting it into the neural network for training.

3.1. Deep neural network — Convolutional Neural Network (CNN)

CNN is a more efficient type of artificial neural network obtained by combining artificial neural networks with deep learning theory. CNN mainly consists of convolutional layers, pooling layers, fully connected layers, and a Softmax layer. The input is the feature map to be detected, followed by convolution and pooling feature extraction operations on the feature map, then flattening and a fully connected layer to generate a one-dimensional vector, finally passing through the Softmax layer to obtain the category probability corresponding to the feature, selecting the maximum probability as the determined type of the feature map.

(1) Convolutional Layer

The convolutional layer's function is to extract certain target features from the input image by sliding the convolutional kernel. Convolution is categorized as shallow convolution and deep convolution; shallow convolution is limited to extracting low-level features such as lines and edges, but deep convolution is capable of extracting more intricate features. Each convolutional layer comprises many kernels, with each element of the kernel representing a weight coefficient and a bias term, analogous to the neurons in a feedforward neural network. The parameters of the kernel include kernel size, stride, and padding, which together determine the size of the feature image output by the kernel. The larger the kernel size, the richer and more complex the input features extracted.

(2) Pooling Layer

After the convolution extracts features, the pooling layer down-samples the feature map obtained from the convolution to achieve selection and dimensionality reduction of the output features. The pooling layer has a predetermined pooling function that replaces the result of individual points in the feature map with statistical information of the features in the neighboring area. Common pooling sizes are 2×2 with a stride of 2, as shown in Figure 4.

2	4
6	5

Figure 4: Feature map after pooling

DOI: 10.54254/2755-2721/121/2025.19842

The prevalent pooling operations are maximal pooling and average pooling. Maximum pooling identifies the highest value among four points, whereas average pooling computes the mean value of the four points, hence minimizing redundant information inside the model.

(3) Fully Connected Layer

Every output node is linked to all input nodes; this configuration is referred to as a completely connected layer. The fully connected layer often resides at the conclusion of the convolutional neural network, functioning as a component of the classifier. This layer can learn high-quality image information features and map the learned features to classification labels, enabling the neural network to classify input data. Every neuron in the fully connected layer is linked to all neurons in the preceding layer, computing the output value by the integration of weights and biases; hence, the layer has a high number of parameters and computational load.

(4) Softmax Layer

Softmax is mainly used in the multi-class classification process. It transforms the output of the last layer of the neural network into a probability distribution, representing the predicted probabilities of different categories, mapping the original output to an interval between 0-1, and the sum of these values is 1, representing the probabilities of each category.

(5) Activation Function

The activation function's purpose is to incorporate non-linear combinations into the neural network, allowing it to adjust to diverse complex application contexts. Numerous activation functions exist, with regularly utilized ones including Sigmoid, ReLU, and Tanh, among others[14].

3.2. Deep neural network - Recurrent Neural Network (RNN)

RNN mainly deals with sequential value data and can handle data that is locally unrelated or data with variable lengths on the time dimension. It consists of an input layer, a hidden layer, and an output layer. The arrows pointing to the feedback loop update in the hidden layer achieve the function of memory in time[15-16].

As shown in Figure 5, depending on the specifics of the problem, the number of hidden units, inputs, and outputs in the recurrent neural networks framework will also be adjusted accordingly[17]. Under the same target task, these adjustments in quantity will also have a corresponding impact on the model's accuracy and convergence speed. If the structure suitable for the target task is multiple inputs corresponding to a single output, one alternative solution is as follows: first use a multi-input multi-output approach, then apply an attention mechanism to the multi-output results, and finally obtain a single output. This replacement method can improve the model's accuracy to a certain extent [18].

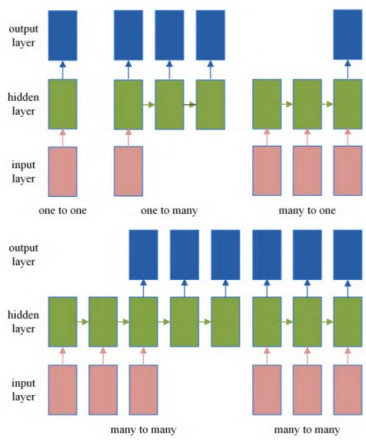


Figure 5: Common stucture of recurrent neural networks

DOI: 10.54254/2755-2721/121/2025.19842

The way that recurrent neural network units are connected has an extremely important impact on the performance of recurrent neural networks. Skip connections can alleviate the problem of vanishing gradients in recurrent neural networks[19]. The two types of recurrent neural networks shown in Figures 6 and 7 have the same number of units and layers, but their different connection methods often lead to different effects[20]. The output result of any intermediate time node in the bidirectional recurrent neural networks shown in Figure 6 depends not only on the inputs before that time node but also on the inputs after that time node[21]. The stacked recurrent neural networks depicted in Figure 7 seeks to enhance model accuracy through increased network depth; however, this simultaneously elevates computational demands and the likelihood of encountering local optima[22].

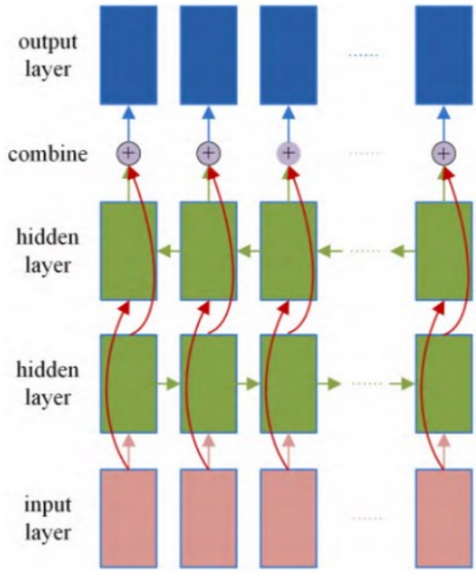


Figure 6: Stucture of bidirectional recurrent neural networks

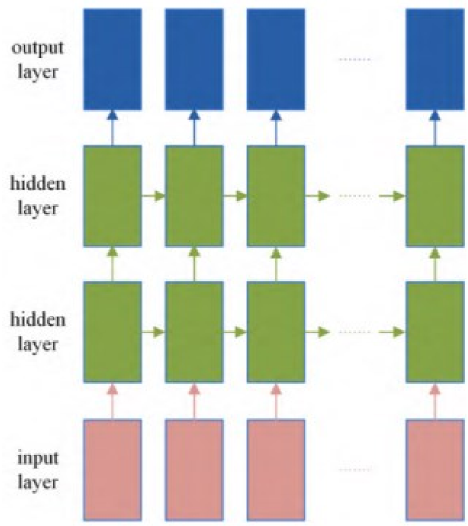


Figure 7: Stucture of stacked recurrent neural networks

4. Future and outlook

Modulation recognition technology has advanced significantly in recent years and attained notable results; nonetheless, certain limitations and challenges persist. Despite the advancement of modulation recognition technology over several decades and significant enhancements in its performance, the introduction of novel communication modulation techniques has rendered classical modulation recognition algorithms, as well as the majority of deep learning-based algorithms, capable only of closed-set recognition for familiar modulation types. This means they can only classify blind signals as one of the known modulation types within the established set. In non-cooperative modulation recognition, once a new modulation method signal appears, the existing closed-set recognition methods will incorrectly identify it as one in the modulation set, rather than judging it as a new signal. Therefore, how to achieve open-set recognition of modulation methods is particularly important. Although some domestic scholars have conducted research on open-set recognition, there is still a long way to go before practical application.

Secondly, currently, most domestic and foreign scholars tend to focus on the research of supervised deep learning algorithms, which rely on a large number of samples with corresponding labels. Nonetheless, a substantial quantity of unlabeled data exists. When the sample size is inadequate, the algorithm's performance will decline significantly. In non-cooperative military communications, it is often difficult to obtain a large number of samples, especially some sensitive military communication signals that are restricted by the geographical environment and enemy confidentiality requirements. Only a very small number of samples can be obtained normally, and during wartime, the battlefield will be filled with a large number of communication signals. It is impossible to adapt to the rapidly changing battlefield situation by collecting these signals one by one and labeling them for recognition. In non-cooperative civilian communications, a large number of unlabeled data can be relatively easily obtained based on diversified signal acquisition techniques and large-capacity storage technologies. However, annotating the collected unlabeled data requires a lot of time and economic costs. Consequently, investigating modulation recognition on unlabeled signals is a valuable area of research.

Finally, many modulation recognition algorithms, in order to improve recognition accuracy, often require a large amount of data and powerful hardware platform computing resources. Especially, deep learning-based modulation recognition methods often require computers equipped with high-performance graphics cards to train neural networks for a long time to achieve better results. However, such algorithms mostly only stay on the simulation stage and are often difficult to deploy. In practical applications, the algorithm needs to be deployed on mobile hardware platforms, not on platforms with strong computing power. Therefore, how to design lightweight (miniaturized, low-power) modulation recognition systems that can be used for actual deployment is of great research significance[23].

5. Conclusion

This paper initially presents modulation recognition technology independent of deep learning, addressing its definition, principles, and pros and cons, specifically the likelihood ratio recognition method rooted in decision theory and the pattern recognition method based on feature extraction. The research begins with deep neural networks and subsequently presents modulation recognition technology with deep learning, emphasizing the examination of convolutional neural networks. Finally, it briefly summarizes the existing problems in modulation recognition technology and proposes possible research directions. AMR still lacks research on open-set recognition and unlabeled samples; future efforts should strengthen studies in these two areas. Additionally, research on lightweight (compact, low-power) modulation recognition systems that can be deployed in practical applications should also be enhanced.

References

[1] *Sang Lei, Sun Mingyang. (2024). Building a Spectrum Security Defense for the Guardians of the Airwaves[N]. Heilongjiang Daily(005). DOI:10.28348/n.cnki.nhjrb.2024.003628.*

[2] *MiR Autonomous Mobile Robot. (2024). Large MiR AMR fleet helps Jianlin realize the three-layer intelligent upgrade of "people", "goods" and "factory"[J]. Automation Expo, 41(08):20-22.*

[3] *Xue Dejin, Shang Tao, Dong Shijun, etc. (2022). Recognition Method of Modulation Mode of Non-cooperative Communication Signal[J]. Modern Defence Technology, 50(05):152-159.*

[4] *Lin Chong, Yan Wenjun, Zhang Limin, Wang Yujia. (2021). Overview of modulation recognition of communication signals. Journal of China Academy of Electronics and Information Technology, 016 (011), 1074-1085.*

[5] *Liang Wei, He Shuyuan. (2018). Mean empirical likelihood method. Advances in Mathematics (China), 47 (2), 9*

[6] *Chen Yuming, Chen Guijing. (2010). Development of GLR test infinite mixture population[J]. Journal of Anhui University (Natural Science Edition), 34(06):7-12.*

[7] *KIM K, POLYDOROS A. (1988). Digital modulation classification: the BPSK versus QPSK case[C]//MILCOM 88, 21st Century Military Communications — What's Possible?'. Conference record. Military Communications Conference. Piscataway, NJ, USA: IEEE, 431-436.*

[8] *LAY N E, POLYDOROS A. (1995). Modulation classification of signals in unknown ISI environments[C]// Proceedings of MILCOM '95. Piscataway, NJ, USA: IEEE, 170- 174.*

[9] *PANAGIOTOU P, ANASTASOPOULOS A, POLYDOROS A. (2000). Likelihood ratio tests for modulation classification[C]//MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority (Cat. No.00CH37155). Piscataway, NJ, USA: IEEE, 670-674.*

[10] *Li Shiping, Chen Fangchao etc. (2012). Modulation identification algorithm based on cyclic spectrum characteristics in multipath channel[J]. Journal of Computer Applications, 32(8): 2123-2127.*

[11] *Wu Peijun, Hou Jin, Lv Zhiliang, Gui Meishu, Zhang Yuexiao, Chen Zeng. (2019). A multi base phase modulation signal recognition algorithm based on constellation diagram recovery. Telecommunication Engineering, 059 (005), 549-555*

[12] *Dou Yazheng. (2024). Implementation Method of Speech Recognition System Based on Deep Learning[J]. Audio Engineering, 48(10):74-76. DOI:10.16311/j.audioe.2024.10.021.*

[13] *Wang Hai, Hang Xiaohu, Han Shuhe etc. (2024). Analysis of Weld Image Recognition Method Based on Deep Learning[J]. Auto Time, (22):142-144.*

[14] *Zang Haiyan, Yan Wemjum, Zhang Limin etc. (2022). An Overview of Communication Signal Modulation Recognition[J]. Journal of Naval Aviation University, 37(01):126-132.*

[15] *Kakalou, I., Psannis, K. E., Krawiec, P., & Badea, R. (2017). Cognitive radio network and network service chaining toward 5G: Challenges and requirements. IEEE communications Magazine, 55(11), 145-151.*

[16] *Chen Changmei, Li Yanbin. (2020). Research on modulation pattern recognition based on convolutional neural network[J]. Information Technology, 44(1): 101-106.*

[17] *Andrej Karpathy blog. May 21, 2015. The Unreasonable Effectiveness of Recurrent Neural Networks. https:// karpathy.github.io/2015/05/21/rnn-effectiveness/*

[18] *Bahdanau, D. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv: 1409.0473.*

[19] *LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, 521(7553), 436-444.*

[20] *Hu Zhongyuan, Xue Yuan, Zha Jiajie. (2023). Survey on Evolutionary Recurrent Neural Networks[J]. Computer Science, 50(03):254-265.*

[21] *Graves, A., Fernández, S., & Schmidhuber, J. (2005, September). Bidirectional LSTM networks for improved phoneme classification and recognition. In International conference on artificial neural networks (pp. 799-804). Berlin, Heidelberg: Springer Berlin Heidelberg.*

[22] *Graves, A., Mohamed, A. R., & Hinton, G. (2013, May). Speech recognition with deep recurrent neural networks. In 2013 IEEE international conference on acoustics, speech and signal processing (pp. 6645-6649). Ieee.*

[23] *Guo Yunxin, Ma Hong. (2018). A review of research on modulation recognition of communication signals based on machine learning [J]. Electronic Measurement Technology, 41 (24): 107-111. DOI: 10.19651/j.cnki-emt.1802322*