Proceedings of CONF-MLA 2025 Symposium: Applied Artificial Intelligence Research
DOI: 10.54254/2755-2721/2025.BJ27246

Comparative Analysis of CNN and ResNet in Automatic Modulation Recognition

Yiliang Su

Northeastern University at Qinhuangdao, Qinhuangdao, China
2054672372@qq.com

Abstract. Automatic modulation recognition (AMR) is a crucial task in wireless communication systems, enabling intelligent signal processing and spectrum management. Among various deep learning methods, Convolutional Neural Networks(CNNs) and Residual Networks(ResNets) consistently outperform traditional approaches in AMC tasks. However, the effectiveness of residual structure in shallow networks is still worth studying. With both networks constrained to four layers, this study conducts a comprehensive experimental study comparing CNN and ResNet models for AMC using the RML2018 dataset, a widely adopted benchmark for modulation classification. Both networks are trained under identical settings (epoch=200, batch_size=128) to ensure a fair comparison. The results reveal that ResNet's classification accuracy degrades to random-guessing levels at -10dB SNR, performing significantly worse than CNN. As the Signal to Interference plus Noise Ratio(SNR) increases to 20dB, the gap between the two gradually narrows, though ResNet still slightly trails behind CNN. However, ResNet demonstrates superiority in high SNR scenarios. This finding contradicts the prevailing assumption that ResNet universally outperforms other architectures in deep learning applications. This paper analyzes potential reasons for this phenomenon, including residual connection design and feature extraction efficiency.

Keywords: Automatic Modulation Recognition, Convolutional Neural Networks, Residual Networks, Signal-to-Noise Ratio, Deep Learning

1. Introduction

With the rapid development of wireless communication technologies, AMR has become a critical component in intelligent spectrum management, cognitive radio, and military signal surveillance [1]. In military communications, AMR can quickly identify the modulation mode of enemy signals and provide key support for electronic countermeasures. In spectrum management, AMR helps to dynamically monitor spectrum usage, optimize spectrum allocation, and improve spectrum utilization. Traditional AMR methods are categorized into two types: likelihood-based and feature-based approaches [2,3]. However, traditional methods are constrained by artificial features and limited model expressiveness, hindering their ability to handle modern wireless communication's complexity and dynamism. The deep neural network automatically extracts signal features through end-to-end learning, which avoids the limitation of traditional methods relying on manual feature

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.BJ27246

design [4]. The powerful nonlinear modeling ability can effectively deal with the distortion problem in communication. In neural networks, ResNet represent an architectural enhancement over CNN, with both sharing core components like convolution and pooling layers. This paper compares the performance of ResNet and CNN in shallow networks(4 layers), and studies the practical utility of the residual mechanism in wireless signal feature extraction.

While CNNs are widely adopted for AMC tasks, ResNet was initially devised to address gradient vanishing in deep networks. However, whether residual connections also work in shallow networks is still worth studying. In this paper, CNN and ResNet models are systematically compared in terms of performance in shallow networks based on a data set, a widely recognized benchmark for modulation classification. Both models are trained under identical hyperparameters (e.g., epoch=200, batch_size=128) to ensure a fair evaluation. Experimental verification demonstrates the residual structure's role in shallow networks and evaluates whether it embodies a universal design principle beyond mitigating degradation in deep networks. Given their reduced parameter requirements and suitability for edge device deployment, studying performance in shallow networks holds significant practical importance for real-world applications.

2. Related work

The application of AMR in signal processing and communication in the past decades has been widely studied, and many researchers have improved the traditional methods based on likelihood estimation and feature-based. Javier Maroto et al. introduced a hybrid approach integrating maximum likelihood label generation with adversarial training, which improves the accuracy and adversarial robustness of AMC by allowing the model to learn "uncertainty" in complex scenarios, and found that this improvement can be effectively transferred to online scenarios [5]. Maqsood H. Shah et al. developed a computationally efficient likelihood ratio-based framework for orthogonal STBC MIMO systems, demonstrating 90% classification accuracy under blind channel conditions even at low SNR levels in the blind channel scenario [6]. Pejman Ghasemzadeh et al. propose a low-complexity framework based on HoS feature extraction and the SVM classifier to optimize the performance of blind AMC in real-time applications [7]. The traditional methods are mainly developed under the assumption of Gaussian noise, and there are still defects in computational complexity and characteristic generalization ability. In the research, under non-Gaussian noise, the classification accuracy of traditional methods is only 0.1-0.6 at low SNR [8].

Recently, deep learning has gradually replaced traditional methods in terms of modulation recognition. It is good at automatically extracting robust features from raw I/Q data, so as to effectively overcome the limitations of artificial feature design [9]. The commonly used DL models include networks such as CNN, ResNet and recurrent neural network(RNN). Saeed Mohsen et al. proposed two CNN models to realize modulation classification for mathematical vectors (RadioML2016.10a) and image data sets, respectively, and verified their high performance through ROC curves [10]. Yang Peng et al. proposed an automatic modulation classification method based on deep residual neural network modeling with a mask, which extracted signal features by complex value and real value ResNet. It outperformed previous approaches under conditions of small sample size and poor signal-to-noise ratio, with its higher performance validated on the simulated data set [11].

DOI: 10.54254/2755-2721/2025.BJ27246

3. Deep learning model

3.1. CNN

CNNs, as a dominant subclass of feedforward neural networks, have become ubiquitous in computer vision tasks ranging from image recognition to visual pattern analysis, establishing themselves as fundamental architectures in modern deep learning systems [12]. As illustrated in Figure 1, the canonical CNN architecture employs a hierarchical cascade of convolutional layers (for local feature extraction), pooling layers (for spatial downsampling), and fully-connected layers (for high-level reasoning), collectively enabling progressive data abstraction through dimensionality reduction while preserving discriminative features for enhanced model generalizability. CNN has made significant advancements in areas such as image recognition, object detection, image generation, and many others.

The convolutional layer is one of the most critical components in CNNs, automatically extracting local features from input data through convolution operations. Beyond simple image filtering, the convolutional layer can learn features at various scales and levels within images or other types of data, providing valuable insights to the model. The formula for the convolution layer is:

$$Y\left(i,j\right)=\left(X*Y\right)\left(i,j\right)=\sum_{m=0}^{k_h-1}\sum_{n=0}^{k_w-1}X\left(i+m,j+n\right)\cdot K\left(m,n\right) \tag{1}$$

Among them, $*$represents the convolution operation, $X(i+m,j+n)$ represents the value of the input at position $(i+m,j+n)$, $K(m,n)$ and represents the weight of the convolution kernel at $(i+m,j+n)$.

The pooling layer performs downsampling to reduce spatial dimensions and computational complexity, with max pooling preserving salient features and average pooling smoothing local responses. Common types of pooling layers include maximum pooling, mean pooling and spatial pyramid pooling.

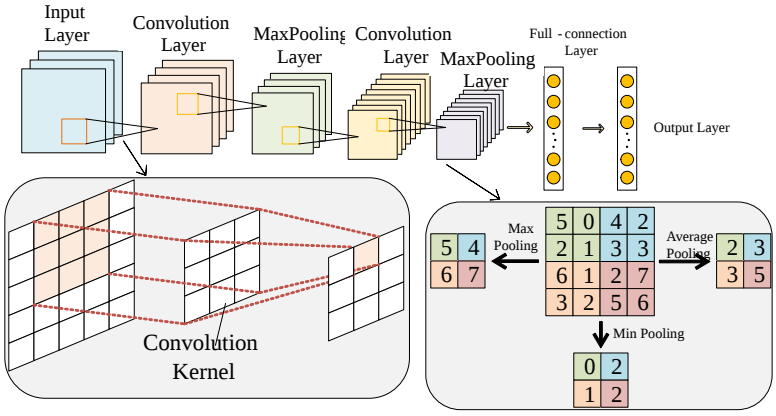

Figure 1. The structure of CNN

3.2. Resnet

ResNet is a special structure of CNN [13]. Empirical studies in deep learning consistently demonstrate that network depth serves as a critical determinant of model representational capacity. With increasing layer depth, the network's receptive field expands exponentially, enabling

DOI: 10.54254/2755-2721/2025.BJ27246

hierarchical feature extraction from low-level edges to high-level semantic patterns—a theoretical advantage formalized by the universal approximation theorem. However, deeper networks often face degradation issues: as the network depth increases, accuracy may plateau or even decline. This is because deep networks can suffer from gradient disappearance or explosion, making it challenging to train deep learning models. ResNet addresses this by introducing 'residual blocks' shown in figure 2, which involve adding skip connections between layers to directly feed the input into the output.

$$\frac{\partial RELU(H(x))}{\partial x} = \frac{\partial RELU}{\partial (F(x)+x)} \frac{\partial (F(x)+x)}{\partial x} = 1 + \frac{\partial F(x)}{\partial x} \tag{2}$$

The gradient of residual units is 1 higher than in ordinary CNN structures due to the backpropagation algorithm formula 2 and chain rule. During back propagation, the gradient can be directly connected to a more forward network layer through the residual connection. Thus, ResNet can prevent the gradient disappearance issue that occurs with deeper networks.

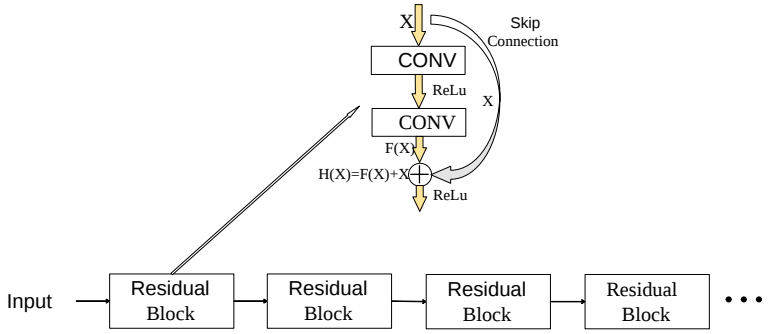

Figure 2. Residual network

4. Experimental and result analysis

4.1. Experiment introduction

This experiment was conducted in a Windows 11 environment, utilizing the deep learning framework TensorFlow and GPU 5070ti to accelerate the training process. The experiments used CNN models and ResNet models to identify digital modulation signals. To ensure the effectiveness of the experiment, the RadioML2018.01A dataset, a widely used modulation classification benchmark, was employed. The dataset includes 24 modulation methods and 26 signal-to-noise ratios(-20, -18,…,28, 30), with an input format of (1024,2), representing a complex signal with 1024 samples. The dataset consists of 2,555,904 data points, with 60% used for training and 20% for testing.

Considering that the computer has only a single GPU and limited performance and training efficiency, to ensure the smooth progress of the experiment, each network was tested under only one SNR condition, and five different SNRs were tested. CNN and ResNet in this paper adopt the network structure shown in Figure 3 respectively. After debugging, the model parameter setting is shown in Table 1.

DOI: 10.54254/2755-2721/2025.BJ27246

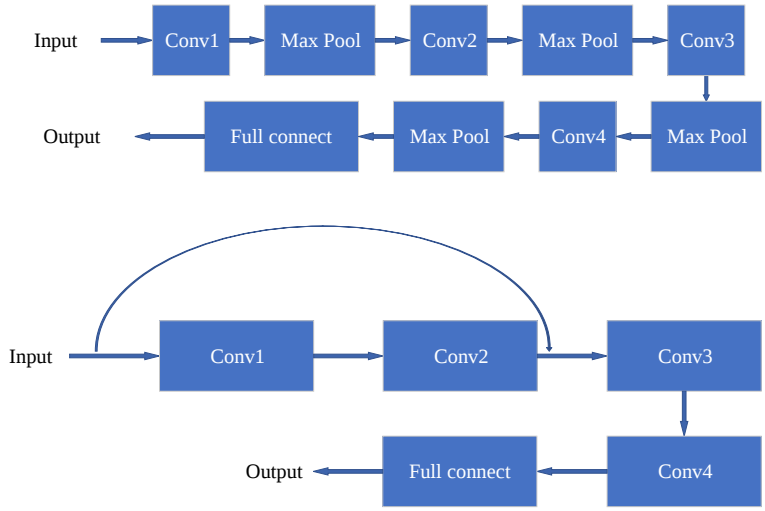

Figure 3. CNN and ResNet structure

Table 1. Model parameter setting of CNN and ResNet

Parameter	CNN	ResNet
Activation Function	ReLU	ReLU
Pooling	Max pooling	Nil
Convolution size	2 X 8	2 X 8
Optimizer	AdamOptimizer	AdamOptimizer
epoch	200	200
btch_size	128	128

4.2. Experimental result

The evaluation protocol employed five strategically selected Signal-to-Noise Ratio (SNR) levels (-10 dB, 0 dB, 10 dB, 20 dB, and 30 dB) spanning the operational range of practical wireless communication systems. By summarizing the experimental results, it is evident that the CNN model's accuracy under these five SNR conditions is shown in Figure 4, while the ResNet model's accuracy is illustrated in Figure 5. The overall accuracy comparison between CNN and ResNet is presented in Table 1. Both architectures exhibit monotonically increasing classification accuracy with higher SNR conditions, though the marginal improvement diminishes significantly beyond 10 dB—a phenomenon consistent with Shannon's channel capacity theory. However, at most SNRs, ResNet's accuracy is 1 to 5 percent lower than CNN's, which is contrary to its results in the deep network.

DOI: 10.54254/2755-2721/2025.BJ27246

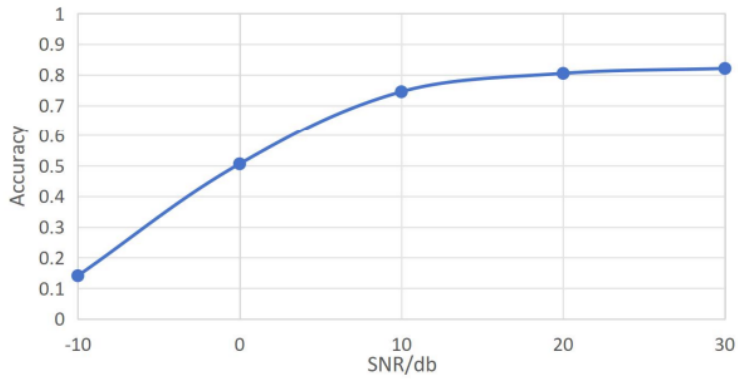

Figure 4. Accuracy of CNN under different SNR conditions

As shown in Figure 6 and Table 2, ResNet performs poorly at extremely low SNR. At an SNR of-10dB, ResNet's accuracy is only 0.04, indicating that it is almost not learning and is essentially guessing randomly. This may be due to the signal interference caused by its residual connections. As previously mentioned, the residual network used in the experiment contains only one residual block, which means that when the noise power far exceeds the signal power, the residual connection directly passes the strong noise to the lower layers, leading to feature confusion. In contrast, the continuous downsampling structure of CNNs can better filter out noise and retain effective features.

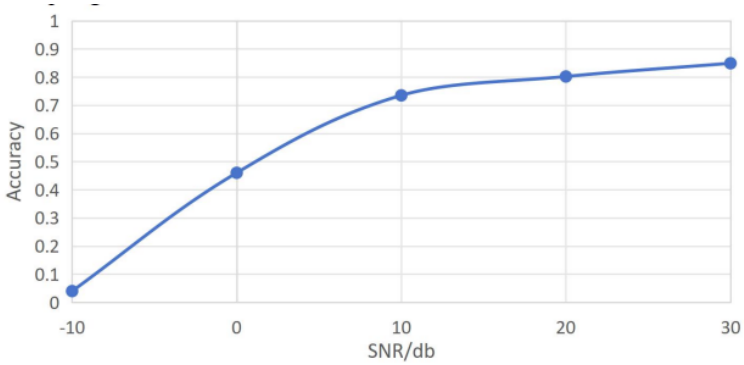

Figure 5. Accuracy of ResNet under different SNR conditions

When the SNR increases to 20db, the gap between the two gradually narrows and almost disappears. The residual blocks of ResNet begin to show a slight advantage, as the residual structure better preserves the original signal features in high SNR conditions, avoiding the loss of details that can occur with multi-layer CNN convolutions. At this point, both architectures achieve usable performance, but ResNet converges faster.

DOI: 10.54254/2755-2721/2025.BJ27246

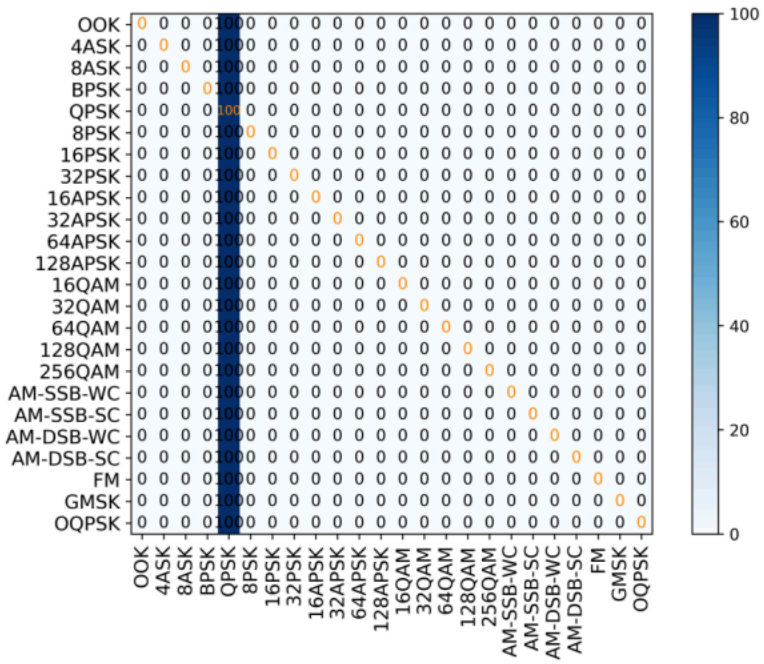

Figure 6. ResNet confusion matrix in SNR=-10db

When the signal quality is above 20db, ResNet shows a trend of exceeding CNN, indicating that when the signal quality is high enough, its local adjustment extraction and cross-layer connection are more effective under clear signal conditions.

Table 2. Accuracy of CNN and ResNet under different SNR conditions

SNR	CNN	ResNet
-10db	0.141	0.040
0db	0.502	0.460
10db	0.742	0.735
20db	0.805	0.802
30db	0.821	0.849

5. Conclusion

Based on the RML2018 data set, this study compares the performance of CNN and ResNet models in AMC tasks under shallow network conditions. The experimental results reveal that CNN consistently outperforms ResNet under low SNR conditions. Specifically, ResNet nearly fails to learn meaningful features at -10 dB, likely due to residual connections amplifying noise when signal quality is poor. Conversely, CNN's layered downsampling structure proves more robust in filtering out noise and preserving useful features. As the SNR improves, the performance gap between the two models narrows, and the minimal difference indicates that their performance is nearly identical. At 20 dB and above, ResNet gradually surpasses CNN, demonstrating faster convergence and slightly higher accuracy. This study indicates that the benefits of ResNet's residual connections are not universally applicable in shallow networks (4 layers) and are significantly influenced by the signal-to-noise ratio. In actual applications, CNNs may be more appropriate for low SNR

DOI: 10.54254/2755-2721/2025.BJ27246

circumstances, but ResNets demonstrate potential in high-quality signal environments. The efficacy of ResNet with more layers warrants further investigation in the future.

Acknowledgment

First, I would like to extend my deepest gratitude to the esteemed professors and teachers. Their invaluable guidance and profound insights have been crucial to my thesis writing. Their deep knowledge, unwavering patience, and constructive criticism have deepened my understanding of the field. I am particularly grateful to Professor Yang for his meticulous guidance, inspiring discussions, and constant encouragement, which have motivated me to overcome challenges and continuously improve my work. I also wish to express my sincerest thanks to all the teaching assistants, who have provided thought-provoking suggestions and consistent support throughout my academic journey. Further, I would like to thank all my friends and parents for their encouragement and support. Without all their enlightening instruction and impressive kindness, I could not have completed my thesis.

References

[1] Lu, H., Zhou, S., Peng, Z., Zhu, H., Zhang, H., & Hu, Q. (2024). Research progress on modulation recognition of radio signals. Journal of Sichuan University of Science & Engineering (Natural Science Edition), 37(6), 31–42.
[2] Abu-Romoh M, Aboutaleb A, Rezki Z. 2018. Automatic modulation classification using moments and likelihood maximization [J]. IEEE Communication Letters, 22(5): 938-941.
[3] Kharbech S, Dayoub I, Zwingelstein-Colin M, et al. 2016. On classifiers for blind feature-based automatic modulation classification over multiple-input-output channels [J]. Iet Communications, 10(7): 790-795.
[4] He, X., & Zeng, W. (2024). End-to-end estimation method based on deep neural network. Digital Communication World, (8), 83–85.
[5] Maroto, J., Bovet, G., & Frossard, P. (2023, June). Maximum likelihood distillation for robust modulation classification. In ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 1-5). IEEE.
[6] Shah, M. H., & Dang, X. Y. (2020). An effective approach for low-complexity maximum likelihood based automatic modulation classification of STBC-MIMO systems. Frontiers of Information Technology & Electronic Engineering, 21(3), 465-475.
[7] Ghasemzadeh, P., Banerjee, S., Hempel, M., & Sharif, H. (2019). Accuracy analysis of feature-based automatic modulation classification with blind modulation detection. In 2019 International conference on computing, networking and communications (ICNC) (pp. 1000-1004). IEEE.
[8] Ma, J., Lin, S. C., Gao, H., & Qiu, T. (2019). Automatic modulation classification under non-Gaussian noise: A deep residual learning approach. In ICC 2019-2019 IEEE International Conference on Communications (ICC) (pp. 1-6). IEEE.
[9] Peng, S., Sun, S., & Yao, Y. D. (2021). A survey of modulation classification using deep learning: Signal representation and data preprocessing. IEEE Transactions on Neural Networks and Learning Systems, 33(12), 7020-7038.
[10] Mohsen, S., Ali, A. M., & Emam, A. (2024). Automatic modulation recognition using CNN deep learning models. Multimedia Tools and Applications, 83(3), 7035-7056.
[11] Peng, Y., Guo, L., Yan, J., Tao, M., Fu, X., Lin, Y., & Gui, G. (2023). Automatic modulation classification using deep residual neural network with masked modeling for wireless communications. Drones, 7(6), 390.
[12] Chua, L. O. (1997). CNN: A vision of complexity. International Journal of Bifurcation and Chaos, 7(10), 2219-2425.
[13] He, F., Liu, T., & Tao, D. (2020). Why resnet works? residuals generalize. IEEE transactions on neural networks and learning systems, 31(12), 5349-5362.