Research on Several Neural Network Structure for Automatic Modulation Recognition

Yidong Xu

{15050865676@163.com}

School of Electronic Information Engineering, Xidian University, Xi'an, Shaanxi, 710126, China

Abstract. With the rapid development of communication technology, Automatic Modulation Recognition (AMR) based on Deep learning (DL) performs well relying on its unique advantages. However, due to the wide variety of neural networks, it is important to compare and analyze their performance and applicability under specific conditions. In this paper, we select convolutional neural network (CNN) and Residual networks (Resnet), and continuously deepen the depth of the residual network to explore the influence of the accumulation of residual blocks. After simulating and analyzing the recognition effects of different network structures under -12 to 30 signal-to-noise ratio (SNR) conditions, the experimental results show that under the experimental conditions set up in this paper, the recognition rate of Resnet is about 4.8% higher than that of CNN on average when SNR is higher than 0db. After accumulating one and two residual blocks and fine-tuning the model to improve the recognition rate, the recognition rate of both networks obtained from the improvement exceeds 90% when SNR is higher than 10db.

Keywords: Automatic modulation recognition, Deep learning, Convolutional neural network, Residual neural network.

1 Introduction

Modulation recognition of communication signals refers to judging the modulation method used in the signal under the premise of unknown modulation information content and modulation parameters, thus providing a basis for the demodulator to correctly select the demodulation algorithm, and ultimately obtaining useful information content. Nowadays, the channel transmission environment becomes more and more complex, resulting in the increasing difficulty of modulation identification. Therefore, the traditional method based on manual identification is becoming more difficult to meet the requirements. Recently, DL has been widely used in the fields of image processing, speech processing and so on [1-2]. Thanks to the strong data analysis ability of DL, more and more scholars choose to use it in automatic modulation recognition. This DL-based AMR can achieve both feature extraction and recognition, thus getting rid of the complex and difficult manual feature. AMR can be categorized into three types according to the difference in the representation of the signal after preprocessing [3]. The first category is to extract expert experience features directly from the received signals and then use neural networks for recognition. The second category is to convert the IQ signal obtained from sampling into the form of an image, thus converting the signal recognition problem into an image recognition problem. The third category is to directly use the in-phase and orthogonal signals as the input to the neural network. The internal learning

CONF-MLA 2024, November 21, Adana, Turkey

DOI 10.4108/eai.21-11-2024.2354628

mechanisms of deep learning networks are not fully transparent since their decisions rely heavily on complex computations of a large number of parameters, that are difficult to interpret intuitively. Many neural network-based automatic feature extraction and classification algorithms also sacrifice a certain degree of interpretability. These factors lead to the so-called “black box” problem. To develop a better performing "black box" and increase the accuracy of automatic modulation detection, we continuously alter the neural network structure in this article, concentrating on deepening the residual network by stacking residual blocks.

2 Literature review

Artificial Neural Networks (ANN) have been used as classifiers for modulation recognition as a result of the development of artificial intelligence, machine learning, and other technologies [4]. As the feature extraction potential of neural networks has been continuously explored in research in fields such as images and natural language, the increasingly strong feature extraction capability makes neural networks more and more effective in recognizing communication modulated signals.

In 2016, T.J.O 'Shea et al. demonstrated for the first time the ability of shallow convolutional neural networks to recognize communication modulated signals. The article uses the RML 2016.10 dataset, which has become a benchmark dataset for training and testing modulation recognition performance [5]. Li et al. proposed a VHF signal modulation identification method based on anti-noise processing and deep sparse filtering CNN model, which first extracts the cyclic spectrum of the sampled signals and represents it with a low rank, and then uses sparse filtering to train the CNN, which improves the robustness of the model [6]. Yang et al. [7] proposed a dual-path modulation recognition mode consisting of improved residual stacks and long short-term memory(LSTM). By introducing transfer learning, the model was transferred from RML2016.10b to RML2018.01a and HisarMod 2019.01. The result of the experiment showed the accuracy consistently exceeded 90% when SNR is high, proving that the model has good robustness.

Chen Changmei et al. trained a network to recognize seven modulation modes by expanding the signal as a two-dimensional constellation diagram as input data. The recognition rate can reach 97. 99% when SNR is higher than 5dB and 100% when SNR is higher than 9dB [8]. Fugang Liu et al. extracted the signal's higher-order accumulation, SNR, instantaneous features, and cyclic spectra, which were then fed into a parallel network of CNN and gated circulation unit (GRU). This method still has an 80% recognition rate at a SNR of -10db for the eight modulations selected in the paper [9]. Kun Liu et al. proposed a bi-directional convolutional selective recursive deep network architecture, which combines three network structures, CNN, GRU and DNN, to realize high-precision modulation recognition. The recognition rate reaches 98.1% at a Doppler frequency of 200Hz when the SNR is 8dB [10]. Han et al. proposed a signal modulation recognition method based on multi-feature fusion and constructed a deep learning network with a two-branch structure to extract the features of IQ signals and multi-channel constellations respectively. This method can construct a more complete representation of signal features, which can help to improve the classification accuracy [11].

Zeng et al. focused on the effects of carrier frequency offset and sampling rate offset on AMR. A novel transformer-based method named TransGroupNet is designed in this paper, which is capable of extracting signal deep features from the instantaneous amplitude, phase and frequency domains. Experimental results show that this model has up to 98% correctness under conditions of high SNR but large offsets [12]. Wei proposes an innovative Multi-Dimensional Shrinkage Block (MDSB) to address the problem of low recognition accuracy of AMR when SNR is low. Using this architecture, the model performs very well on four public datasets, including RML2016.01a, RML2016.01b, RML2018.01a and HisarMod2019.01 [13]. Research in recent years has shown that there are a variety of network framework structures for AMR based on DL, which have shown great improvement in terms of time reduction and accuracy improvement.

3 Algorithm introduction

This paper utilizes CNN and Resnet to achieve end-to-end signal modulation type identification by automatically learning and extracting features from signal I/Q data without human intervention.

3.1 CNN-based methods

Convolutional neural network is one of the mainstream neural network structures. This network shows excellent ability in processing data with spatial features and has been widely used in image processing, speech recognition and target recognition. In this network, the input layer accepts image data, and the convolutional and pooling layers extract the most representative features to reduce the risk of overfitting and enhance network generalization. In 2012, Hinton proposed the Dropout algorithm. This algorithm refers to discarding some neurons in the iterative process, so as to reduce the model’s dependence on certain local features and improve the model’s generalization ability. Due to the outstanding spatial feature extraction capability of CNN, an attempt has been made to introduce it into the study of AMR [14].

For this experiment, the selected CNN comprises four convolutional layers, four pooling layers, and two fully connected layers [15]. To mitigate overfitting, a dropout layer is incorporated into the network architecture. The number of filters in the convolutional layers is decreasing, which is beneficial to reduce the training time and get better performance. Figure 1 illustrates the network's architecture, while Table 1 provides a comprehensive list of its specific parameters.

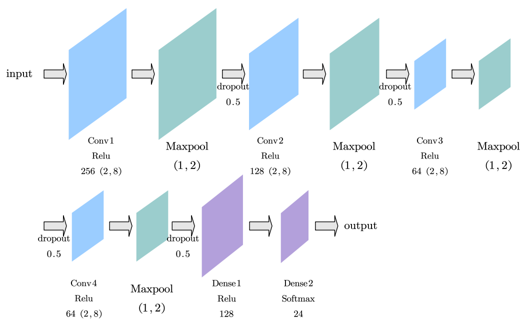


Fig. 1. Structure of CNN

Table 1. Table of CNN parameters

Layer	Output dimensions
input	2*1024
Conv1	2*1024*256
Max_pool1	2*512*256
Dropout1	2*512*256
Conv2	2*512*128
Max_pool2	2*256*128
Dropout2	2*256*128
Conv3	2*256*64
Max_pool3	2*128*64
Dropout3	2*128*64
Conv4	2*128*64
Max_pool4	2*64*64
Dropout4	2*64*64
Flatten	8192
Dense1	128
Dense2	24
Trainable Par.	1,777,304

3.2 Resnet-based methods

Generally speaking, as the network width and depth increase, its performance will improve. However, networks with too many layers may lead to gradient vanishing or gradient explosion. In order to avoid this phenomenon as much as possible, Kaiming He proposed residual neural networks [16].

Resnet is constructed based on CNN, which inherits the basic components of CNN. The core of this network is the introduction of residual connectivity, which passes the results of the previous layer of learning directly to the next layer. This architecture enables Resnet to preserve all the informational content of the signal while effectively utilizing the deep and shallow information of the data. The error does not change drastically due to the deepening of the network layers. The structure of the residual block is shown in Figure 2.

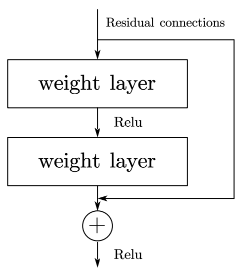


Fig. 2. Standard structure of the residual block [17]

The residual neural network chosen in this paper has four convolutional layers, two fully connected layers, and uses the dropout algorithm to prevent overfitting. Its structure is shown in Figure 3 [17].

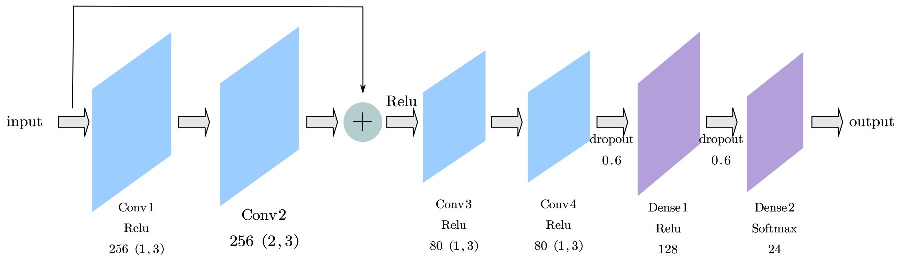


Fig. 3. Structure of Resnet

Based on the established network architecture, this paper seeks to enhance the recognition accuracy of the network by stacking additional residual blocks. For this purpose, the following two network structures are proposed in this paper.

The first network proposed in this paper (Proposed1) adds a residual block to the original network, which is the unit consisting of Conv3 and Conv4 in Figure 4. After many trials, this paper chooses to set the size of both convolution kernels to 2*3 to get better results. The network structure is shown in Figure 4.

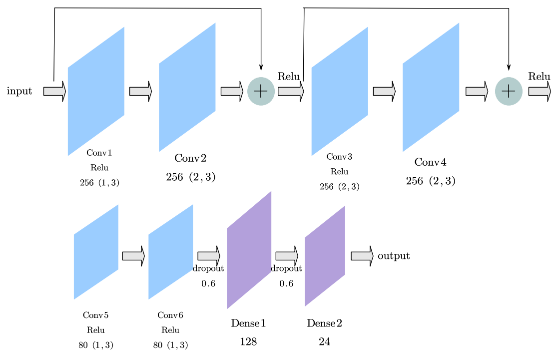


Fig. 4. Structure of Proposed 1

The second proposed network (Proposed2) enhances the architecture by incorporating three residual blocks while decreasing the number of convolution kernels to 128 in each of the initial six layers. This modification aims to simultaneously reduce the parameter count and improve recognition performance. Figure 5 illustrates the structure of this network.

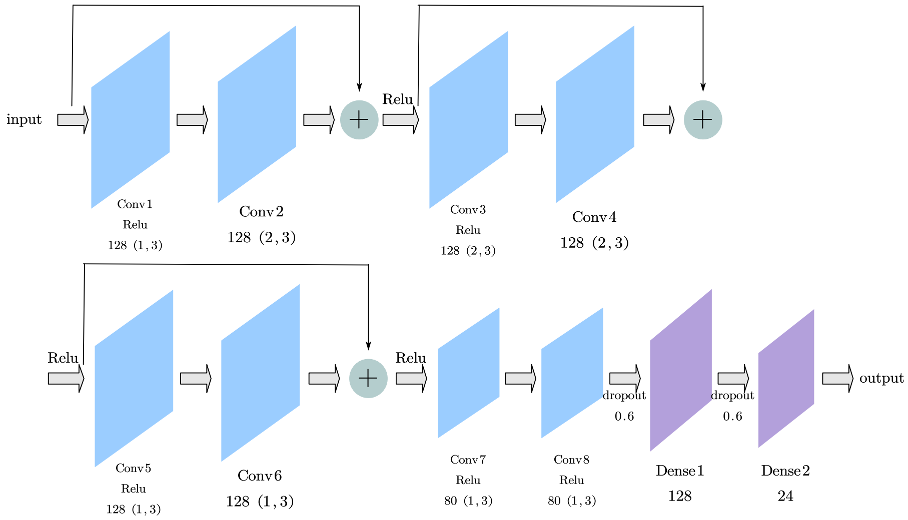


Fig. 5. Structure of Proposed 2

In order to clearly compare the structure of the three kinds of Resnet, the parameters of Resnet, Proposed1 and Proposed 2 are all listed in Table 2.

Table 2. Resnet, Proposed1, Proposed2 network parameter table

Layer	Output dimensions		
	Resnet	Proposed1	Proposed2
Input	2*1024	2*1024	2*1024
Conv1	2*1024*256	2*1024*256	2*1024*128
Conv2	2*1024*256	2*1024*256	2*1024*128
Add	2*1024*256	2*1024*256	2*1024*128
Conv3	2*1024*80	2*1024*256	2*1024*128
Conv4	2*1024*80	2*1024*256	2*1024*128
Add	/	2*1024*256	2*1024*128
Conv5	/	2*1024*80	2*1024*128
Conv6	/	2*1024*80	2*1024*128
Add	/	/	2*1024*128
Conv7	/	/	2*1024*80
Conv8	/	/	2*1024*80
Dropout1	2*1024*80	2*1024*80	2*1024*80
Flatten	163840	163840	163840
Dense1	128	128	128
Dropout2	128	128	128
Dense2	24	24	24
Trainable Par.	21,450,040	22,236,984	21,419,192

4 4 Simulation

4.1 Introduction to the dataset

This paper uses the RadioML2018.01A dataset, which was generated by O'Shea using GNU radios in a relatively good real-life laboratory environment, and has been widely used for the training and testing of AMR models [18]. This dataset contains a total of 24 modulation modes. The SNR of each modulation signal ranges from -20db to 30db, with a step size of 2db. This dataset consists of three kinds of parameters, which are I/Q signal, modulation mode, and SNR. During the experiment, this paper chooses to split the dataset according to SNR to observe the performance of the model under different SNRs. This paper selects SNR ranging from -12db to 30db, 60% of the dataset is used for training the model; 20% is used for model verification but not for training, and the remaining 20% is used to test the trained model.

4.2 Experimental platform and Parameter settings

This experiment uses an NVIDIA GeForce RTX 3070 LapTop GPU with 8GB of GPU memory. The system environment for the experimental simulation is CUDA 11.2 and cuDNN 8.1, and the deep learning environment is tensorflow-gpu 2.10 and keras 2.10.

The model training uses the adam optimizer, the loss function is categorical cross entropy, and the callback function is used to dynamically adjust the learning rate. The initial learning rate is 0.001. If the validation loss does not improve within 5 epochs, the learning rate will be multiplied by 0.5. The number of epochs is set to 200. The training process will stop if the validation loss fails to show improvement over a span of 50 epochs. The model which performs best on the validation dataset is retained and evaluated using the test dataset.

4.3 Simulation results

The simulation test results are shown in Figure 6. Propose 1 is the first improvement based on the original Resnet, and Proposed 2 is the second improvement based on the original Resnet.

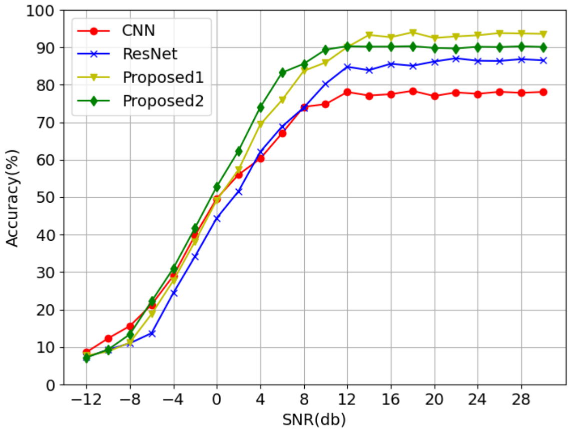
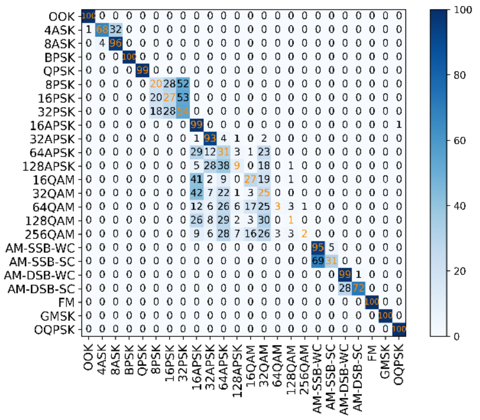


Fig. 6. Comparison of recognition accuracy of four networks under different SNRs

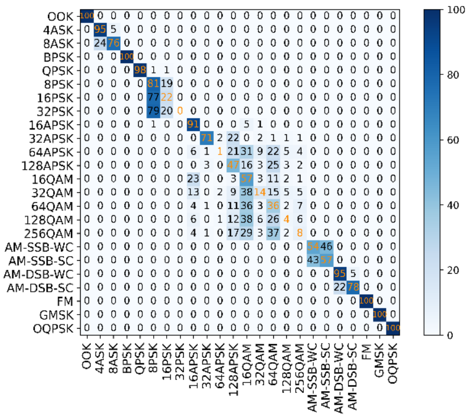
As can be seen from Figure 6, for the original CNN network and Resnet, when SNR is higher than 8db, the advantage of the Resnet becomes obvious. When SNR is greater than 10db, the average recognition accuracy of Resnet is 8.13% higher than that of the CNN.

The two improved networks obtained in this paper based on the existing Resnet have significantly improved the recognition rate. When SNR is greater than 10db, the recognition of both networks reaches 90%. When SNR is 0db to 12db, Proposed2 performs best, while when SNR is greater than 12db, Proposed1 performs best, with an average recognition accuracy of 93.30%, which is about 3.21% higher than Proposed2.

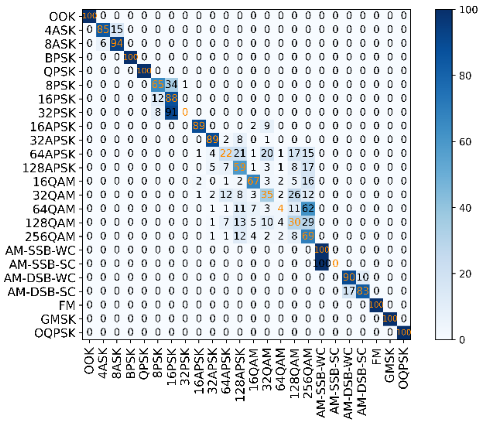
It can be seen that in this experiment, the deeper Proposed2 performs better than Proposed1 at a lower SNR, but when SNR is high enough, the advantage of Proposed1 is greater. In order more intuitively compare and analyze the recognition capabilities of the four models in this paper for signals of various modulation modes, this paper chooses to use the form of confusion matrix to show the test results of the four models when SNR is 4dB, 10dB and 18dB, as shown in Figure 7.



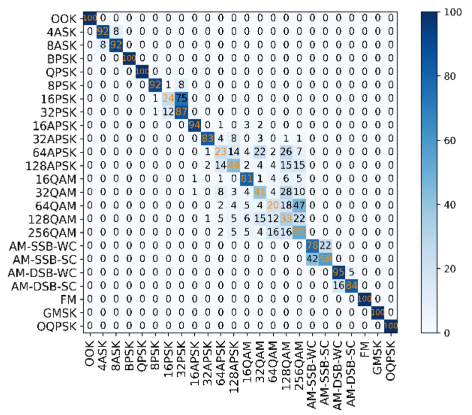
CNN, 4db



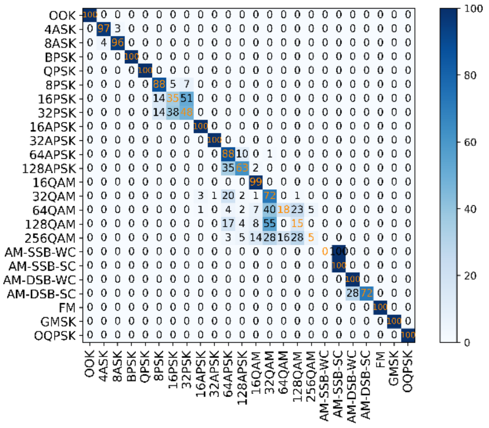
Resnet, 4db



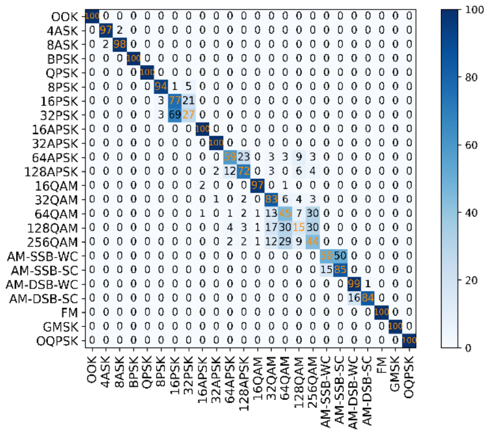
Proposed 1, 4db



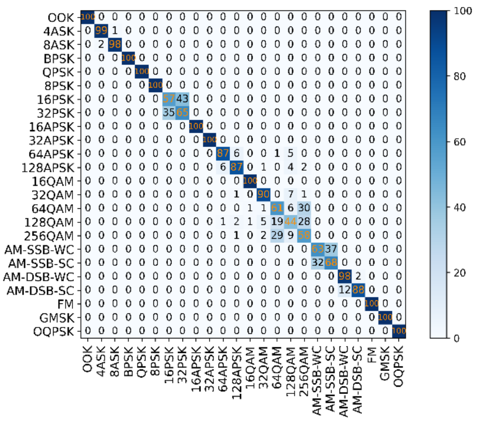
Proposed 2, 4db



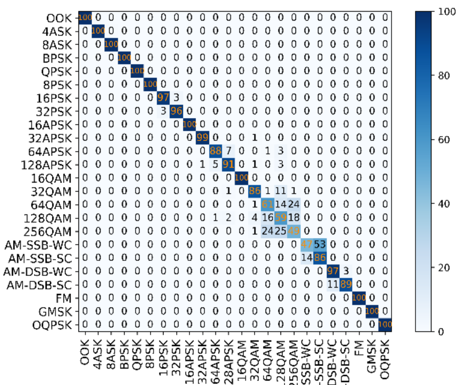
CNN, 10db



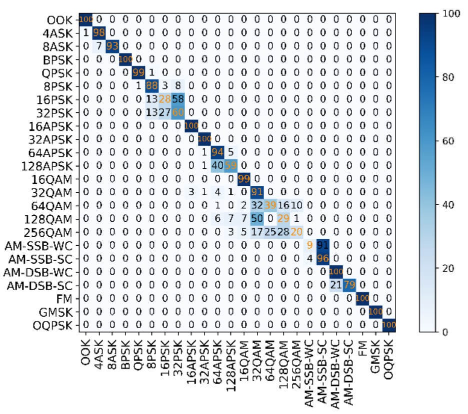
Resnet, 10db



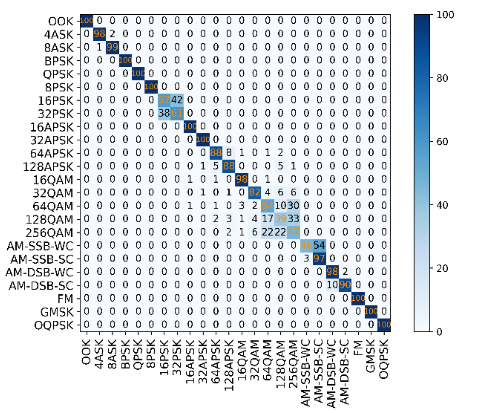
Proposed 1, 10db



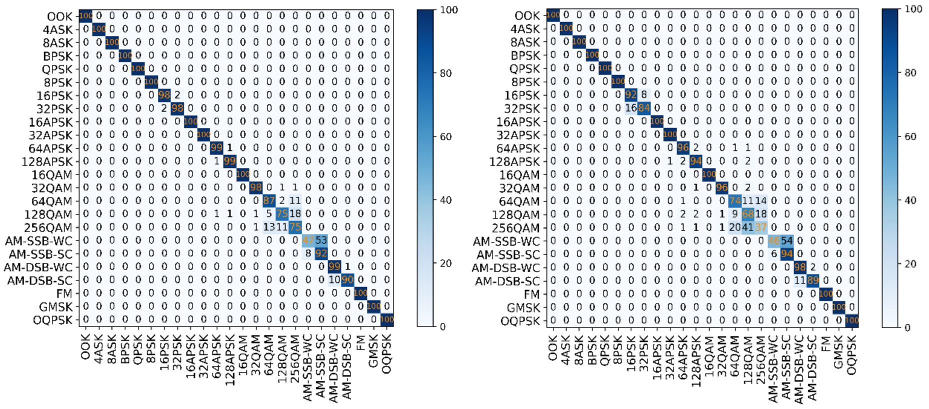
Proposed 2, 10db



CNN, 18db



Resnet, 18db



Proposed 1, 18db Proposed 2, 18db

Fig. 7. Confusion matrix results of four network modulation recognitions with SNR of 4db, 10db and 18db

As can be seen from Figure 7, when SNR is 4 dB, the modulation modes with the worst recognition effect of the four networks are all QAM. But overall, the color blocks of the confusion matrices of the two improved networks are closer to the diagonal position than before the improvement, reflecting that the improvement is effective. When SNR is 10db, the recognition results of Resnet and CNN are not much different. The poorly recognized modulation methods are PSK, QAM and AM-SSB. CNN recognizes AM-SSB-WC and AM-SSB-SC as AM-SSB-SC, while Resnet basically recognizes 16PSK and 32PSK as 16PSK. These two networks can basically not recognize QAM. Proposed 1 is a big improvement over the previous two networks. It does not completely fail to distinguish between two modulation modes, but the recognition of confused modulation modes is basically the same as that of the original CNN and Resnet. The biggest advantage of Proposed 2 over the other three networks is that it can basically completely distinguish PSK, and it also performs better in distinguishing different QAMs and different AM-SSBs. When SNR is 18dB, CNN still cannot distinguish AM-SSB-SC, while Resnet is still slightly better than CNN. For Purposed1 and Purposed2, the modulation modes that are confused in their identification are mainly QAM and AM-SSB, but Proposed1 performs better in identifying QAM than Purposed 2.

The structures of CNN and Resnet used in this experiment are derived from the existing solutions [15][17]. The experiment in this paper verifies these two models. Then this paper stacks residual blocks on the basis of the original Resnet to obtain Proposed1 and Proposed 2. The number of their parameters is basically the same as the original Resnet, but the recognition effect is improved. However, this method is data-driven and requires a large amount of labeled data. Compared with the research method based on small samples, this paper still has shortcomings. In the future, we can try to use new algorithms that only require a small number of samples to complete the training process.

5 Conclusion

Based on the public dataset RML2018.01a, this paper selects CNN and Resnet to simulate and analyze the effect of modulation recognition, and studied the impact of the stacking of residual blocks on Resnet in the AMR task. The results show that Resnet, based on the CNN structure can achieve great performance improvements by accumulating residual blocks. Propose1 and Propose 2 have a recognition accuracy of 90% when SNR is 10db, which is a level that the original CNN and Resnet cannot achieve under high SNR conditions. Moreover, as the network depth increases, the recognition rate of Resnet also improves to a certain extent when SNR is low. However, the performance of Proposed1 and Proposed2 in distinguishing 128QAM from 256QAM and AM-SSB-WC from AM-SSB-SC is worse than in other cases, which indicates that these two networks still have room for further improvement in this regard.

References

[1] Boopathi S, Kanike U K. 2023. Applications of Artificial Intelligent and Machine Learning Techniques in Image Processing[M]//Handbook of Research on Thrust Technologies' Effect on Image Processing. IGI Global, 151-173.

[2] Triantafyllopoulos A, Schuller B W, İymen G, et al. 2023. An overview of affective speech synthesis and conversion in the deep learning era[J]. Proceedings of the IEEE, 111(10): 1355-1381.

[3] Zhang Qianqian, Wang Yu, Lin Yun, et al. Review of automatic modulation recognition methods based on deep learning[J]. Radio Communication Technology, 2022, 48(04): 697-710.

[4] Zhang Chengchang, Yu Sa, Xu Yu, et al. 2022. A review of the application of artificial neural networks in modulation recognition[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 34(02): 181-192.

[5] O'Shea T J, Corgan J, Clancy T C. 2016. Convolutional radio modulation recognition networks[C]//Engineering Applications of Neural Networks: 17th International Conference, EANN 2016, Aberdeen, UK, September 2-5, 2016, Proceedings 17. Springer International Publishing, 213-226.

[6] LI R, LI L, YANG S, et al. 2018. Robust Automated VHF Modulation Recognition Based on Deep Convolutional Neural Networks [J]. IEEE Communications Letters, 22(5): 946-949.

[7] Huogen Y, Lingzhu Z, Guangxue Y, et al. 2021. IRLNet: A Short-Time and Robust Architecture for Automatic Modulation Recognition[J]. IEEE ACCESS, 9143661-143676.

[8] Chen Changmei, Li Yanbin. 2020. Research on modulation pattern recognition based on convolutional neural network[J]. Information Technology, 44(01): 101-106.

[9] Liu F, Zhang Z, Zhou R. 2021. Automatic modulation recognition based on CNN and GRU[J]. Tsinghua Science and Technology, 27(2): 422-431.

[10] LIU K, XIANG X, LIANG Y, et al. 2021. Automatic Modulation Recognition through Wireless Sensor Networks in Aero-nautical Wireless Channel [J]. IEEE Sensors Journal, 21(20): 23125-23132.

[11] Han H, Yi Z, Zhu Z, et al. 2023. Automatic Modulation Recognition Based on Deep-Learning Features Fusion of Signal and Constellation Diagram[J]. Electronics, 12(3): 552.

[12] R. Zeng, Z. Lu, X. Zhang, J. Wang and J. Wang, 2024. "Convolutional Neural Network Assisted Transformer for Automatic Modulation Recognition Under Large CFOs and SROs," in IEEE Signal Processing Letters, vol. 31, pp. 741-745, doi: 10.1109/LSP.2024.3372770.

[13] T. Wei, Z. Li, D. Bi, Z. Shao and J. Gao, 2023. "Adaptive Multi-Dimensional Shrinkage Block for Automatic Modulation Recognition," in IEEE Communications Letters, vol. 27, no. 11, pp. 2968-2972, Nov. doi: 10.1109/LCOMM.2023.3314623.

[14] Zhang, Fuxin, et al. 2022. "Deep learning based automatic modulation recognition: Models, datasets, and challenges." Digital Signal Processing 129: 103650.

[15] K. Tekbıyık, A. R. Ekti, A. Görçin, G. K. Kurt and C. Keçeci, 2020. "Robust and Fast Automatic Modulation Classification with CNN under Multipath Fading Channels," 2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring), Antwerp, Belgium, pp. 1-6.

[16] K. He, X. Zhang, S. Ren, et al. 2016. Deep Residual Learning for Image Recognition[C]. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 770-778.

[17] X. Liu, D. Yang and A. E. Gamal, 2017. "Deep neural network architectures for modulation classification," 2017 51st Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, pp. 915-919, doi: 10.1109/ACSSC.2017.8335483.

[18] O'SHEA T J, ROY T, CLANCY T C. 2018. Over-the-Air Deep Learning Based Radio Signal Classification [J]. IEEE Journal of Selected Topics in Signal Processing, 12 (1): 168-179.