Advances in Computer, Signals and Systems (2025)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2025.090120
ISSN 2371-8838 Vol. 9 Num. 1

Classification of RF Signals Based on Image and Sequence Inputs

Jingren Wang, Yaxuan Xu, Jing Zhang

School of Information and Intelligent Engineering, Sanya College, Sanya, Hainan, China

Keywords: RF signal classification; ResNet18; sequence input; image input; deep learning

Abstract: Radio frequency (RF) signals are widely used in commercial and military wireless communications, and the accurate classification of such signals is of great theoretical significance and practical application value. This study aims to solve the practical problem of RF signal classification based on a 5-classified signal dataset (abbreviated as 2021 dataset) provided by a company. In this paper, we design ResNet18 based on constellation map input, ResNet18 based on hybrid map input and CNN network based on sequence input, and compare and analyze the performance of the deep learning algorithms by exploring the deep learning algorithms under two input modes: image and sequence. It is shown that the classification accuracy of ResNet18 using hybrid graph input reaches 95.79%, while the CNN model with sequence input performs better in terms of classification accuracy and real-time performance, with an accuracy of 98.22%, and the number of parameters is only about 1/8 of that of ResNet18.

1. Research Background

RF signals are widely used in many fields such as commercial and military wireless communications. Accurate classification of RF signals is important for improving the efficiency of communication systems, optimizing resource allocation, and ensuring information security. Classification of RF signals in high signal-to-noise ratio environments is an important research direction in this field, and the development of deep learning provides new ideas and methods to solve this problem. This study focuses on exploring the application of deep learning algorithms based on image and sequence inputs in high SNR RF signal classification, aiming at comparing the performance of the models with different inputs, and finding more suitable classification models for practical applications.

Automatic Modulation Classification/Recognition (AMC/AMR) [1], as a key link in communication systems, has gone through several stages of development. Traditional automatic modulation recognition methods mainly include maximum likelihood (LB)-based and feature extraction (FB)-based methods. The LB method is actually a multivariate hypothesis testing problem, which realizes the classification of modulation signals with the help of decision criteria [2]. Although better results can be obtained theoretically, the computational complexity is high and the practical application is limited; the FB method achieves better performance in modulation recognition, but faces the problems of complex feature extraction, dependence of classification effect on a priori information, and poor generalization ability.

As deep learning has achieved remarkable results in many fields, its application in RF signal modulation recognition has gradually attracted attention. The modulation recognition method based on convolutional neural network [3] can be divided into image input recognition and sequence input recognition according to the form of input data. Among them, signal modulation recognition using image input already has many better results, in 2019 Peng et al. proposed a modulation recognition algorithm based on constellation diagrams, and utilized AlexNet[3] and GoogLeNet[4] to classify the signals, the results are better and faster than traditional machine learning algorithms[5].Ozturk et al. in 2020 converted signals into time-series and spectral maps to build a CNN model to classify UAV signals, and experiments proved that spectral maps are more advantageous for classification [6]. Considering that there may be information loss when using one image alone, H. Elyousseph et al. in 2021 merged the IQ signal time-series map with the spectrogram according to the three channels of the image and proved that the use of hybrid maps has better results than using only a single image [7]. In terms of sequence input recognition, O'Shea et al. proposed in 2016 to directly input IQ data into a convolutional network for classification, and released a series of radio signal simulation datasets for deep learning, among which RadioML2018.01.A (RML2018) has a rich variety of modulation types and data volumes [8]. team proposed in a subsequent work that adding residual structures to convolutional networks [9] can achieve a high correct classification rate [10]. In conclusion, there have been many deep learning based works in the field of high signal-to-noise ratio signal classification that have obtained better results. In this paper, we refer to these excellent works, and at the same time, combined with the problems of high model complexity and poor generalization that still exist in the existing research, we make further research and optimization based on the self-built signal dataset.

2. Introduction of the dataset

In the process of signal transmission, the transmitted signal is a low-frequency signal that is not suitable for long-distance transmission, so it is necessary to add a high-frequency carrier signal to the original signal to make it suitable for transmission, and this process is called signal modulation. Where the original signal is called modulated signal. The received signal $x(t)$ can be expressed as:

$$x(t) = s(t) + n(t) \tag{1}$$

Where $s(t)$ is the effective post-modulated signal and $n(t)$ is the additive Gaussian white noise which is uncorrelated with the signal [11]. The signal-to-noise ratio (SNR) refers to the ratio of signal to noise and is calculated as follows:

$$SNR = 10lg\frac{P_s}{P_n} \tag{2}$$

Where P_s is the power of the signal and P_n is the power of the noise. The larger the signal-to-noise ratio, the smaller the noise; the smaller the signal-to-noise ratio, the larger the noise. In this paper, we use IQ data to represent the signal, the IQ data is divided into two ways, one way is in-phase component I (in-phase), and the other way is quadrature component Q (Quadrature), their mathematical meanings are as follows:

$$\begin{aligned} I(t) &= a(t)\cos(\varphi(t)) \\ Q(t) &= a(t)\sin(\varphi(t)) \end{aligned} \tag{3}$$

All signal data in this paper are discrete IQ data. The mathematical definitions of I and Q components are based on orthogonal signal principles [12].

2.1 Description of the dataset

The dataset used in this study is the 2021 dataset provided by the partner company. The sampling rate of this dataset is 100kHz, and each segment contains I and Q data, which are analog signal data passing through the transmitter and receiver. There are 28 segments of IQ data in the 2021 dataset, which cover five different modulation categories (refer to Table 1), namely ASK, BPSK, FSK, QAM, and QPSK, among which the categories QAM and QPSK have 8 segments, and the rest of the categories have 4 segments. categories have 4 segments of data.

Table 1 Categories in the 2021 dataset

Frequency Band	1-4	5-8	9-12	13-20	21-28
modulation category	ASK	BPSK	FSK	QAM	QPSK

Table 2 Signal-to-noise ratio in the 2021 dataset

Frequency Band	4N+1	4N+2	4N+3	4N+4
SNR	2	6	10	14

Each category has a different distribution of data at different signal-to-noise ratios. Referring to Table 2, $N = 0, \ldots, 6$. Combining the above two tables, it can be seen that the three categories ASK, BPSK and FSK have one segment of data for each of the four SNRs, while QAM and QPSK have two segments of data for each SNR.

2.2 Data Preprocessing

Since the original data is long and inconvenient to process, this paper refers to the data length of the public dataset RML2018.01. Perform non-recombining cropping on each segment of the original data in steps of 1024.In order to ensure the randomness of the samples, random interval is used for sample cropping.

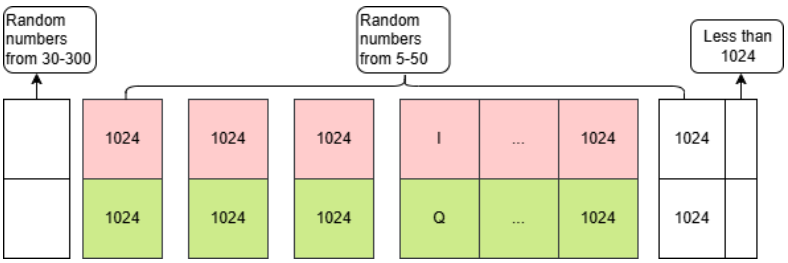

Fig 1 Schematic diagram of random interval cropping

As shown in Fig.1 above, the pink data block is the I-channel data, the green data block is the Q-channel data, and the I-channel data and Q-channel data are cropped and sampled at the same time. At the beginning of each data segment, a random number of 30 to 300 is used as the interval, then the data is sampled in steps of 1024, and then the sampling interval is any random number from 5 to 50 to ensure the randomness of the data. The above operation is repeated until the final length of the remaining data is less than the sum of 1024 and the random number of 5～50. After cropping, the dataset is randomly divided into training set, validation set and test set according to 8:1.5:0.5.The data volumes of the five modulation categories at each signal-to-noise ratio are 3800, 3800, 7602, 7600, 7600, 7600, respectively. The data volumes of the training set, validation set and test set are 72966, 13681, and 4561, respectively.

3. Deep Learning Algorithms Based on Images

3.1 Constellation Diagram

A constellation diagram is a way to visualize the distribution of signal data points. In this study, a scatter plot is drawn with data I as the horizontal axis and data Q as the vertical axis to form a constellation diagram (refer to Fig.2). Signals with different modulation types show different distribution characteristics on the constellation diagram, which provide important visual information for subsequent model training based on the inputs of the constellation diagram, and help the model to learn and distinguish between different modulation types.

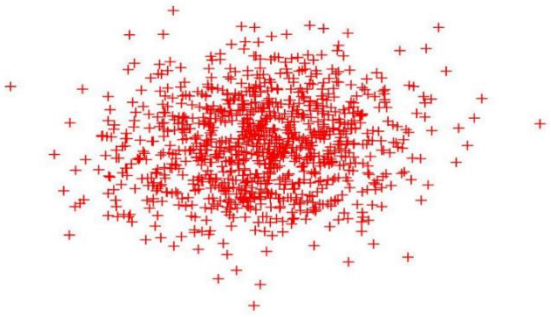

Fig 2 Example of constellation diagram

3.2 Hybrid Map

Hybrid map is an innovative form of image input proposed by Elyousseph et al [7]. It consists of several single-channel images combined into a three-channel image, with the first and second channels being the time-domain maps of the I-channel data and Q-channel data, respectively, and the third channel being the power spectrum obtained by converting the IQ data. The power spectrum reflects the variation of the signal power with frequency, this study uses the welch spectrum estimation method to calculate the power spectrum [13], the specific process includes the steps of input, frame-splitting, windowing, finding the power spectrum of each frame, averaging the power spectrum, and unit conversion (see Fig. 3).

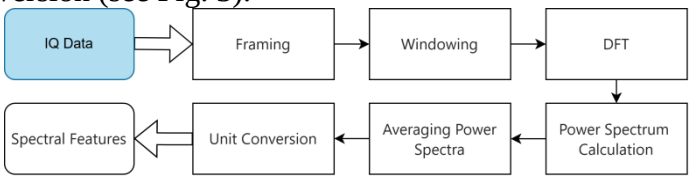

Fig. 3welch spectrum estimation method flowchart

The hybrid diagram combines time domain and frequency domain information, which theoretically provides richer features for the model and helps to improve the classification effect (refer to Fig. 4).

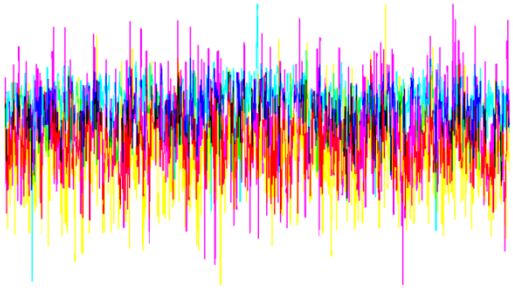

Fig. 4 Example diagram of hybrid map

3.3 ResNet18

The ResNet18 network is a classical deep learning network with good feature extraction capability. It consists of two kinds of Residual block, and different Residual block structures differ in terms of the step size of the convolutional layer, etc. Through the combination of these structures, ResNet18 is able to learn the features of the image effectively. The structure of ResNet18 is shown in Fig.5.

In this study, the constellation map or hybrid map is adjusted to the image size of 224×224×3 and then input to ResNet18 for experiments. During the experiment, 50 rounds are trained and the optimal model on the validation set is saved for subsequent comparison of the performance of the model on the test set under different input methods.

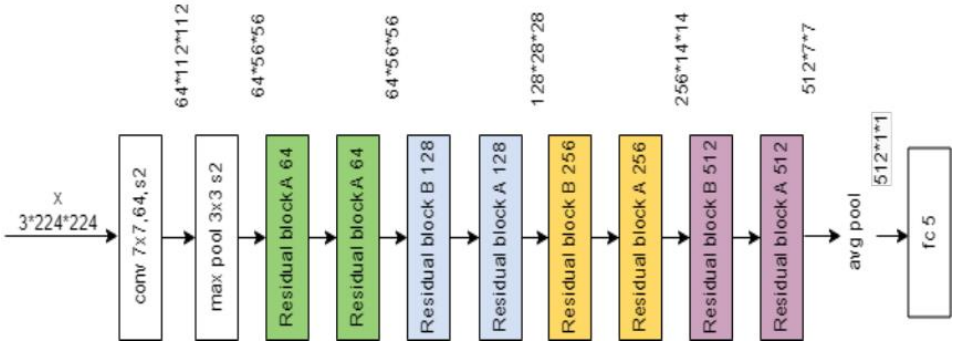

Fig.5 Schematic diagram of ResNet18 network structure

3.4 Numerical Experiments and Analysis

The experimental results in Table 3 show that the hybrid graph input is better than the constellation graph input in terms of overall classification results. In Table 3, HYACC is the predicted correct rate for the hybrid chart input and XACC is the predicted correct rate for the constellation chart input. The table gives the overall correctness of the two models with the correctness under each category. From the table it is clear that there is a significant improvement in the hybrid map comparison with the constellation map. Only the category QAM has 3.31% lower correct rate of classification as compared to the constellation map input, but the correct rate of classification for QPSK is improved by 29.64%.

Table 3 Performance of the two image input methods on test set 1

	Total Correct Rate	ASK	BPSK	FSK	QAM	QPSK
XACC	89.10%	89.18%	97.43%	99.61%	91.30%	57.73%
HYACC	95.79%	100.00%	100.00%	100.00%	87.99%	87.37%

Table 4 Percentage increase in correctness of hybrid map inputs at four signal-to-noise ratios and compared to the correctness of constellation map inputs

(All units are percentages)

Category / SNR	ASK	BPSK	FSK	QAM	QPSK
2	100(+21.67)	100(+11.17)	100(+1.54)	58.5(-34.58)	65.7(+65.2)
6	100(+12.99)	100(+0)	100(+0)	91.8(+5.98)	83.4(+57.4)
10	100(+6.21)	100(+0)	100(+0)	99.48(+8.9)	98.2(+3.24)
14	100(+1.71)	100(+0)	100(+0)	100(+5)	100(+0)

Observation of Table 4 reveals that the classification correctness of the four modulation categories

ASK, BPSK, FSK, and QPSK using the hybrid map as input is better than that of the constellation map input, and the hybrid map input significantly improves the classification effect at low signal-to-noise ratios, especially on the ASK, BPSK, and FSK categories. However, there are preferences for different input features for different categories, for example, the classification correctness of QAM at 2dB is higher than that of hybrid map input for constellation map input.

On the 2021 dataset, the use of hybrid input combining time and frequency domains can well improve the problem of poor classification at low SNR due to noise, but there is a non-negligible problem in converting the signal into image domain and then inputting it into the network: it takes a lot of time for both preprocessing (converting the image) and inference for model training. It takes 49.5 ms to convert single signal data to constellation maps, and 388.3 ms to convert to hybrid maps. Therefore, converting the signal into an image input network has the problem of time-consuming preprocessing and model training inference, which can not simultaneously meet the two requirements of improving classification accuracy and reducing processing time, which affects the real-time nature of the model in practical applications, and is not conducive to the rapid processing of a large number of RF signals.

4. Sequence-based deep learning algorithm

4.1 CNN Model Construction

According to the design idea of VGG image recognition model, combined with the characteristics of RF signals, this study proposes a CNN model for signal recognition (refer to Fig.6).

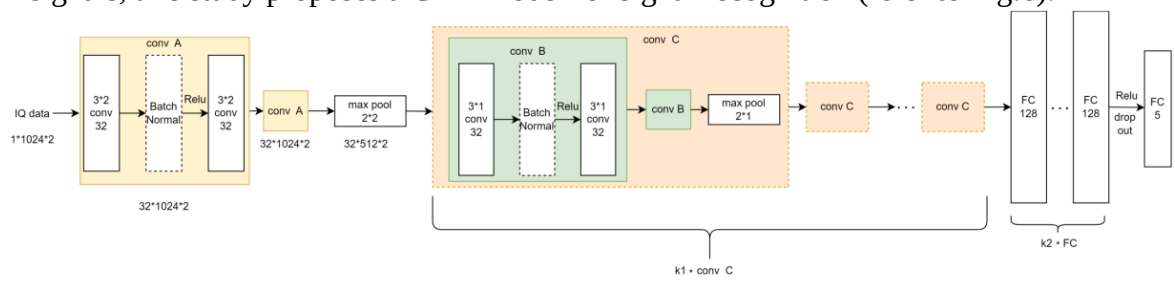

Fig. 6 Schematic diagram of ResNet18 network structure

The model takes the IQ data with a length of 1024×2 as input, and first passes through two convolutional structures named Conv A, whose main purpose is to fuse the data of the two paths of IQ in accordance with the time points; then the output features are compressed to half by Max pool to reduce the data dimension; then it passes through k1 Conv C convolutional structure, which is mainly composed of Conv B convolutional structure and Max This structure is mainly composed of Conv B convolutional structure and Max pooling, which is used to extract the features of the data; after that, it is connected to k2 fully-connected layers with a length of 128, and finally it is connected to the fully-connected layer with a classification number of 5 to output the classification results. In order to prevent overfitting, a dropout layer is placed between each fully connected layer, and the drop rate is set to 0.5.

In order to find the optimal combination of hyperparameters for the CNN model, a series of experiments were conducted in this study. When exploring the role of the BN layer, it is found that the addition of the BN layer can significantly improve the correct rate of the validation set and effectively avoid the gradient explosion problem. When searching for the optimal combination of network layers, it is found that the model works best when the number of fully connected layers is 1 and the number of modules in Conv C is 7. At the same time, the Batch Size also has an important effect on the model effect, with the increase of the Batch Size, the correct rate of the validation set

increases gradually, and reaches the relative optimality when the Batch Size is 1024, and our optimal model is named RML-CNN here.

4.2 Model generalisation performance test

The optimal model obtained using the above hyperparameters is tested on the same source test set 1 and different source test set 2 of the 2021 dataset to explore the generalisation ability of the model, and the training test is conducted on the RML2018 dataset to examine the generalisation ability of the model.

Table 5 Correctness of test sets

	Total Correct Rate	ASK	BPSK	FSK	QAM	QPSK
Correct Rate	98.22%	100.00%	100.00%	100.00%	95.86%	93.63%

The results in Table 5 show that the model has a higher correct rate of 98.22% on the test set, which is not much different from the validation set's correct rate of 98.74%, indicating that the model has good generalisation performance on the 2021 dataset. On the whole, the model achieves 100% correct classification rate in ASK, BPSK and FSK, and compared with the 95.79% correct rate of the model using hybrid graph as image input, the correct rate of the CNN model based on sequence input on the test set has a significant improvement, especially in the low signal-to-noise ratio for the classification of the two types of data, QAM and QPSK, and the parameter count and The number of parameters and computation amount are much smaller than that of ResNet18 used for the image input model, which is more suitable for practical applications. It is worth noting that the number of parameters of the RML-CNN model is 123845, and the computation amount of one data inference is 51427982, which is much smaller than the number of parameters of the ResNet18 model (more than 10 million), and is more suitable for practical applications. The percentage increase in correctness of sequence inputs at four signal-to-noise ratios compared to hybrid map inputs is shown in Table 6

Table 6 Percentage increase in input correctness of sequences with four SNRs and comparison of hybrid maps.(all units are percentages)

Category / SNR	ASK	BPSK	FSK	QAM	QPSK
2	100(+0)	100(+0)	100(+0)	83.6(+25.2)	75.5(+9.8)
6	100(+0)	100(+0)	100(+0)	99.5(+7.7)	98.5(+15.1)
10	100(+0)	100(+0)	100(+0)	100(+0.52)	100(+1.8)
14	100(+0)	100(+0)	100(+0)	100(+0)	100(+0)

However, on the complex Rml2018 dataset, the model has a low classification correctness (refer to Table 7), especially the signal-to-noise ratio of -2dB performs poorly, of course this type of data is not in our 2021 dataset, but this can also side-step the generalisation of the model still needs to be improved, and needs to be further optimised when faced with more complex signal classification tasks.

Table 7 Performance of RML-CNN on RML2018

SNR	-2	2	6	10	14	Total Correct Rate
RML-CNN	27.52%	48.37%	63.47%	67.75%	67.98%	55.03%

5. Conclusion

In this study, two deep learning methods based on image and sequence inputs are explored in depth

for the RF signal classification problem. Among the methods based on image input, the classification effect of hybrid graph has certain advantages in low SNR, but there is the problem of time-consuming preprocessing and model training inference; the CNN model based on sequence input achieves a high classification correctness rate on the 2021 dataset through hyper-parameter optimisation and performs better in real-time, which is more suitable for practical application scenarios. However, the performance of this model on complex datasets needs to be further improved, especially for low signal-to-noise ratios. Future research can consider combining the self-attention mechanism to further optimise the model structure, and exploring more effective feature extraction methods to improve the classification performance of the model in complex environments.

References

[1] Weaver C S, Cole C A, Krumland R B, et al. The Automatic Classification Of Modulation Types By Pattern Recognition [R]. Stanford Univ Calif Stanford Electronics Labs, 1969.

[2] Panagiotou P, Anastasopoulos A, Polydoros A. Likelihood ratio tests for modulation classification[C].MILCOM 2000 Proceedings. 21st Century Military Communications. Architectures and Technologies for Information Superiority (Cat. No. 00CH37155). IEEE, 2000, 2: 670-674.

[3] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.

[4] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2015: 1-9.

[5] Peng S, Jiang H, Wang H, et al. Modulation classification based on signal constellation diagrams and deep learning[J]. IEEE transactions on neural networks and learning systems, 2018, 30(3): 718-727.

[6] Ozturk E, Erden F, Guvenc I. RF-Based Low-SNR Classification of UAVs Using Convolutional Neural Networks[J]. International Telecommunication Union, 2021(5):39-52.

[7] Elyousseph H, Altamimi M L. Deep learning radio frequency signal classification with hybrid images[C]//2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE, 2021: 7-11.

[8] O 'Shea T J, Corgan J, Clancy T C. 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, 2016: 213-226.

[9] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[10] O Shea T J, Roy T, Clancy T C. Over-the-air deep learning based radio signal classification[J]. IEEE Journal of Selected Topics in Signal Processing, 2018, 12(1): 168-179.

[11] Zhang Xiaohong, Wang Lijuan, Ren Shujie. Fundamentals of Digital Signal Processing [M]. Tsinghua University Press, 2007.

[12] Lyons R. A quadrature signals tutorial: Complex, but not complicated[J]. DSPRelated. com| DSP, 2013.

[13] Welch P. The use of fast Fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms [J]. IEEE Transactions on audio and electroacoustics, 1967, 15(2): 70-73.