ISBN: 978-981-18-7950-0
2023 the 13th International Workshop on Computer Science and Engineering (WCSE 2023)
doi: 10.18178/wcse.2023.06.033

A UAV Spectrum Classification Framework Based on ResNet50 Algorithm

Qihang Liu [1], Shuo Su [2], Yao Yao [1], Bing Gui [1] and Zhiyuan Shen[1] +

[1] College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China
[2] College of Electronic and Information Engineering/College of Integrated Circuits, Nanjing University of Aeronautics and Astronautics, Nanjing, China

Abstract. Unmanned Aerial Vehicle (UAV) category recognition is a significant problem for airport airspace operation. An accurate and efficient identification of UAV is expected to fast predict of UAV behavior and implement the defense. This study presents an UAV spectrum classification framework using radio frequency (RF) signals. In the framework, the spectrum is combined with the ResNet50 neural network model. An open source dataset namely DroneRF is used to test the efficiency of the proposed method. The experimental results show that the highest accuracy rate of the proposed framework reached to 100%. Compared with the state-of-the-art method, this study has a higher improvement of accuracy rates.

Keywords: UAV, spectrum detection, artificial intelligence, classification

1. Introduction

As an emerging high-tech product, UAV (Unmanned Aerial Vehicle) has become a current research hotspot. With its advantages like fast speed and high flexibility, UAV is widely used in various fields. The good performance of UAV is applicable to the military needs of land, sea and air, furthermore playing a vital role in civil aviation activities. However, there are some disadvantages that the training conditions is deficient, military and civil aviation safety are threatened. In order to prevent the harm caused by the illegal use and improper operation of UAV, a complete, efficient and accurate UAV identification system has become an urgent need at present.

To facilitate the study of UAV RF signals, Al-Sa'd et al (2019) collected and sorted out different kinds of UAV RF signal data in different flight modes, established a large open source data set DroneRF, and used DNN to classify the signal data. Medaiyese et al.(2021), Kilic et al.(2022) and other research teams used the classical classification algorithms XGBoost, SVM, etc. However the overall correct rate was not high. al-Emadi and Al-Senaid (2020), Nida Kumbasar et al.(2022), Yang Xiaowei et al.(2023), Qian Zhuotao et al.(2021) and other teams tried to extract data features using neural networks with different structures. Among them, Al-Emadi and Al-Senaid (2020) applied 1D-CNN framework to transform one-dimensional data into an image recognition problem, and the classification result of this method was remarkably better than the classical classification algorithm. Nida Kumbasar et al.(2022) introduced HMFFNet model on this basis and changed to a more complex VGG19 neural network structure to extract image features and then use SVM to classify the image features, the correct rate of its method is almost 100%, but the correct rate is low when the segment is 100, and the model training result is poor.

In order to ensure the normal and safe operation of the airport surface area, an efficient and accurate UAV classification and pattern identification framework is necessary. In this study, STFT(Short-time Fourier Transform) data processing method and ResNet50 neural network are used to build a classification and identification framework for UAV RF signals. The composition has high precision and speed, in addition to, it solves the problem of gradient disappearance and explosion.

+ Corresponding author. Tel.: +86-13951916587; fax: +86-25-84893461 .
E-mail address: shenzy@nuaa.edu.cn .

The structure of the paper is followed. Section 2 introduces the basic principle of signal spectrum processing. In Section 3, it constructs the identification framework based on ResNet50 neural network. The comparative experimental are implemented in Section 4. Conclusion is drawn in Section 5.

2. UAV Spectrum Analysis

In this section, the UAV time domain RF signal is converted into a frequency domain signal and represented by a 2D image. For this reason, this study obtained the spectrum map and persistent spectrum of UAV radio frequency signals in the DroneRF data set, and after combining the high frequency and low frequency, the two spectrum graphs were extracted and classified by the classification network. The representation details of the spectrum map and the persistent spectrum are as follows.

2.1. Spectrogram

Short-time Fourier Transform (STFT), its main content is to formulate a pseudo-balance window function to move to generate a stationary signal, so as to calculate the power spectrum at different times[9], its main principle is as follows(1):

$$X(k,m) = \sum_{n=0}^{N-1} w[n]x[n+mR]e^{-j2\pi kn/N} \tag{1}$$

Formula $w[n]$ represents the window function[10]. The value ranges of the variations m and k are $0,1,2\ldots,M-1$ and $0,1,2\ldots,K-1$. In this study, the Kaiser window is used to process each signal. Meanwhile, the signal is divided into equal length and overlapping segments. The short-time Fourier transform of each segment is calculated according to the above equation and all segments are stitched together to form a matrix, and the power of each spectrogram was displayed in decibel step by step, besides, the power distribution is depicted as an image with size-related color mapping. Spectrograms after STFT transformation as show in Fig. 1.

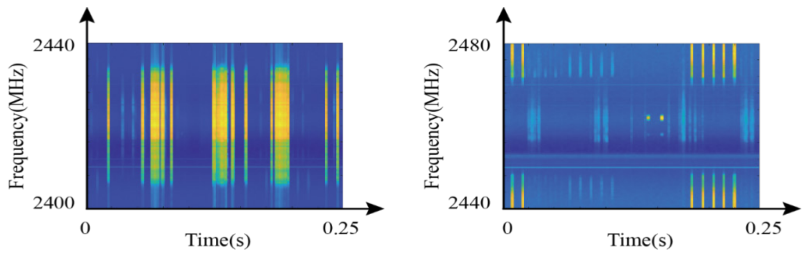

Fig. 1. The left picture shows the low band spectrum, and the right picture shows the high band spectrum

2.2. Persistence spectrogram

The persistence spectrogram of a signal is a time-frequency map that shows the percentage of time that a given frequency appears in the signal, and the longer a particular frequency lasts in the signal, the higher the percentage of time, the brighter the color displayed in the image. The methods for the persistence spectrum is out forward in [11]: using the formula to calculate the spectrum, the power and frequency value is divided into two-dimensional bin. For each moment, a binary histogram of the log of the power spectrum is calculated, adding the matrix element corresponding to each power-frequency bin of the signal energy at that time, and summing the histogram for all moments. Finally, cumulative histograms are plotted against power and frequency, with colors proportional to the logarithm of histogram counts and expressed as normalized percentage. The persistence spectrogram obtained by the above method is shown in Fig. 2.

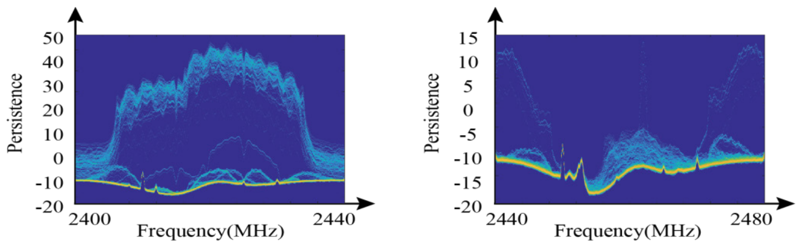

Fig. 2. The left image shows the low band persistence spectrum, and the right image shows the high band persistence spectrum

3. ResNet50-Based Classification Algorithm

This part describes the basic principle and experimental usage of ResNet50 network structure used in detail. The neural network not only provides high quality conditions with its excellent accuracy and speed, but also can solve the problem of gradient vanishing and explosion by cumulative ResNet Blocks. The identification model presented in this study is collectively divided into two parts. As shown in Fig. 3, first, the temporal signal data in DroneRF dataset is transformed into frequency spectrum, transforming signal recognition into image recognition. Second, input part of images into the neural network model to train image features, and then input the other images into the network for classification and recognition.

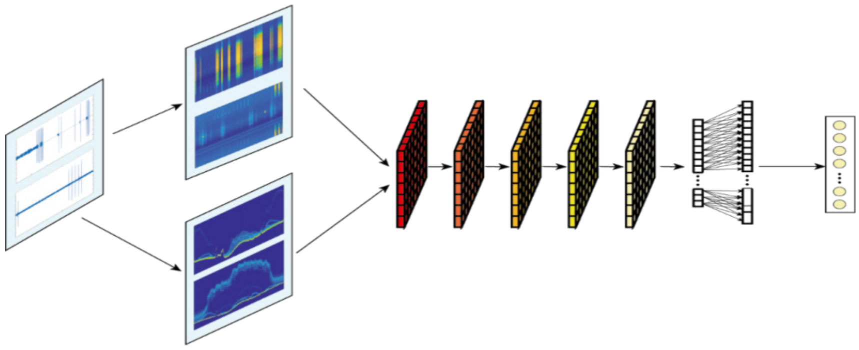

Fig. 3. Model of drone classification

Before the emergence of ResNet, some classical neural networks were subject to gradient disappearance and gradient explosion, which makes the increase of the network layer, the learning ability of the network may decline, as a consequence, network is more difficult to train. Kaiming He et al.(2015) propose Deep Residual Learning's method for the field of image recognition and defined the residual function[12]，as follows(2):

$$F(x) := H(x) - x \tag{2}$$

where x is the input value of each layer and $H(x)$ is the output value after mapping each layer. Then the original mapping function becomes $H(x) = F(x) + x$, as shown in Fig. 4.

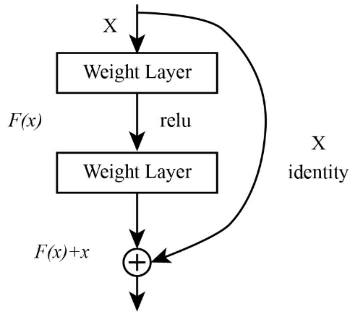

Fig. 4. A building block[12]

Kaiming He et al.(2015) constructing multiple network layers as blocks showing in the following equation. Image vectors can be connected by jumps, which not only speeds up network learning, but reduces deviation in network learning.

$$y = F(x, \{W_i\}) + x \tag{3}$$

x and y are separately representing the input and output vectors of these network layers; the function $F(x, \{W_i\})$ represents the residual mapping that need to be learned.

Kaiming He et al.(2015) obtained the training results of the 1000-layer network by experiment . And three models, ResNet50, ResNet101, and ResNet152, were obtained too. Moreover, ResNet101 obtained a higher correct rate than VGG16 in different datasets.

This study uses the ResNet50 neural network model with one fully connected layer and five stacking blocks, as shown in Fig. 5. The stacked blocks contain $[3, 4, 6, 3]$ convolutional layer stacks, each of them with 3 convolutional layers of different convolutional kernels.

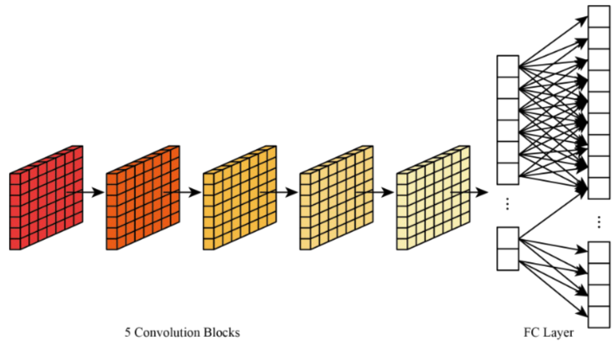

Fig. 5. ResNet50 Layers Settings

4. Case Study

In this section, at the beginning, the evaluation criteria for the performance of the proposed method are specified, then the experimental platform and tools used for the study are presented, after that the experimental method with adjusted parameter settings is given to illustrate how the performance of the proposed UAV detection method is improved. Finally, experimental results are derived and the obtained data and comparison tables are presented.

4.1. Parameter setting

The DroneRF data set was used in this study, an RF based dataset of UAVs operating in different modes, including three different UAVs (AR, Bebop and Phantom) in four modes: off, on and connected, hovering, flying, video recording, and also includes recording segments of RF background activity without UAVs. The dataset contains 227 recording segments (Table 1), for a total of 454 files. Each flight mode recorded RF background activity for 10.25 seconds and RF UAV communication time for approximately 5.25 seconds, as shown in Table 1. The DroneRF dataset stores the RF data in high and low frequency bands, as shown in Table 1.

Table 1. DroneRF dataset content

Drone Type	Segment	Samples	Ratio
AR	81	1620×10^6	35.68%
Bebop	84	1680×10^6	37.00%
Phantom	21	420×10^6	9.25%
No Drone	41	820×10^6	18.06%

The experimental platform are Windows 10 operating system AMD R7 5800H, Intel i7-11800H central processor and hardware device NVIDIA GeForce RTX3060. The data pre-processing part of the experiments and the neural network model was built using Python, Tensorflow 2.11.

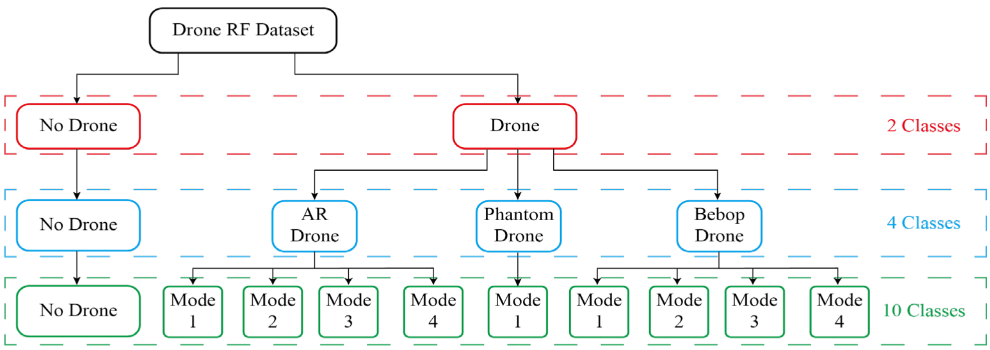

Fig. 6. Classes of Experiments

According to the data classification within the DroneRF data set and the experimental methods of other authors in similar articles, the experiments in this study will follow the classification criteria of dichotomous, quadruple, and decile classes as shown in Fig. 6. In the first instance ,the dataset can be classified into two categories {Drone, No Drone} to determine whether there is a drone in this signal detection area; secondly, the dataset can be classified into four categories {AR Drone, Bebop Drone, Phantom Drone, No Drone} to classify the detected drones; Finally, the dataset can be divided into ten categories {AR Drone Mode 1, AR Drone Mode 2, AR Drone Mode 3, AR Drone Mode 4, ..., Phantom Drone Mode 1, No Drone} to identify the current mode of the UAV.

Each duration recording signals of 250ms in the DroneRF dataset now is currently split into segments for training in some studies, with the number of segments varying from 1-370 segments[1], [2], [7]. In this study, according to the research method in [7], the dataset was split into 1, 10, and 100 segments for training, corresponding to 250ms, 25ms, and 2.5ms, respectively.

In the training, the training and test sets were randomly divided which in order to minimize the risk of obtaining random results and prevent the problem of overfitting, and in this study, 80% of the data was used for training, and the remaining data was used for testing, and the final accuracy derived was the average accuracy of the test set.

4.2. Results and analysis

In this subsection, the experimental results of the recognition model and the comparison with the data results of other articles in the same field are displayed. This paper, along with other articles using the DroneRF data set to evaluate the recognition framework proposed in this paper.

To evaluate the performance of the adopted method, the study used accuracy and F1 score test metrics. They are based on true positive (TP), false positive (FP), false negative (FN), true negative (TN).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Recall = \frac{TP}{TP + FN} \tag{6}$$

$$F1 \text{ score} = 2\frac{Precision \cdot Recall}{Precision + Recall} \tag{7}$$

Table 2. Accuracy of experiments

Figure	Classes	Segment		
		1	10	100
spectrogram	2-Class	**100%**	**100%**	**100%**
	4-Class	97.78%	97.78%	97.78%
	10-Class	86.67%	78.89%	73.33%
persistence	2-Class	**100%**	**100%**	**100%**
	4-Class	96.67%	**100%**	**100%**
	10-Class	88.89%	87.78%	**95.55%**

Table 3. F1 score of experiments

Figure	Classes	Segment		
		1	10	100
spectrogram	2-Class	93.75%	**100%**	**100%**
	4-Class	**100%**	87.93%	89.60%
	10-Class	95.86%	97.59%	79.43%
persistence	2-Class	**100%**	96.77%	**100%**
	4-Class	99.13%	**100%**	**100%**
	10-Class	95.95%	97.59%	**98.18%**

A total of 18 groups of experiments were conducted, and multiple experiments were performed for {spectrogram, persistence spectrogram}, {1, 10, 100}, {dual classification, quadruple classification, decile class} and the average of the experimental results (Accuracy / F1 score) were counted as shown in Table 2, Table 3. Among the 9 sets of comparison experiments, 4 sets of experiments resulted in a higher correct rate using persistence than spectrogram (44.44%), 1 set of experiments resulted in the opposite (11.11%), and the remaining 4 sets of experiments resulted in a 100% correct rate (44.44%).

Other articles have also conducted classification recognition experiments using the DroneRF data set and other methods and frameworks, which are compared here with the experimental results of the same type of articles,, as shown in Table 4. The value N/A means that the corresponding results are not mentioned in the article.

Table 4. Comparison of the result of studies using DroneRF dataset

Literature	Accuracy(%)			F1 score(%)			Number of segment
	2-Class	4-Class	10-Class	2-Class	4-Class	10-Class	
Al-Sa’d et al.(2019)	99.70	84.50	46.80	99.50	78.80	43.00	100
Al-Emadi and Al-Senaid (2020)	99.80	85.80	59.20	99.70	84.60	55.10	100
Medaiyese et al. (2020)	99.96	90.73	70.09	N/A	N/A	N/A	N/A
Allahham et al. (2020)	100	94.60	87.40	100	91.00	77.00	370
Swinney and Woods(2021)	N/A	N/A	91.00	N/A	N/A	91.20	N/A
Nemer et al. (2021)	N/A	N/A	99.20	N/A	N/A	99.10	30
Kiliç et al. (2022)	100	98.67	95.15	100	98.70	94.72	1
Nida Kumbasar et al. (2022)	100	99.55	97.75	100	99.69	97.57	N/A
XiaoWei Yang et al.(2023)	99.8	91.1	70.3	99.7	92.9	66.9	N/A
The proposed ResNet50	100	100	95.55	100	100	98.18	N/A

In fact, not all data points in RF signals can help the network learn. There will be some pixels in the spectrogram that are meaningless for image classification learning; Similarly, as the number of network layers increases in traditional convolutional neural networks, some convolutional layers are not helpful for the learning of classification tasks. Therefore, the introduction of the concepts of "residuals" and " blocks" can effectively alleviate the above phenomenon. Due to the existence of "residuals", the layers pay more attention to the change of input values and output values, and after learning from multi-layer networks, it will be more inclined to reduce the difference between input value and output value, so that the empirical knowledge obtained by network learning tends to be stable. The establishment of "blocks" can accelerate the learning speed of the network, and can quickly skip some layers that are useless for classification tasks.

5. Conclusion

This study built an efficient recognition framework for UAV RF signals. The spectrogram of UAV were transferred to the images for ResNet50 classification. From the experimental results, it was obtained that the persistent spectrum achieved a higher correct rate. When the number of image segments was set as 100, the correct rate of framework recognition is higher, and the correct rates of dichotomous, quadruple, and decile categories were 100%, 100%, and 95.55%, respectively, which are higher than those of competing results 100%, 94.51%, 94.44%. In future studies, more UAV RF signal data may be collected and used for the training of this framework.

6. Acknowledgements

The authors thank the Innovation Training Project of Nanjing University of Aeronautics and Astronautics(202310287052Z).

7. References

[1] M. F. Al-Sa’d, A. Al-Ali, A. Mohamed, T. Khattab, and A. Erbad, “RF-based drone detection and identification using deep learning approaches: An initiative towards a large open source drone database,” *Future Gener. Comput. Syst.*, vol. 100, pp. 86–97, Nov. 2019, doi: 10.1016/j.future.2019.05.007.

[2] Sara Al-Emadi and Felwa Al-Senaid, “Drone Detection Approach Based on Radio-Frequency Using Convolutional Neural Network,” in *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Feb. 2020, pp. 29–34. doi: 10.1109/ICIoT48696.2020.9089489.

[3] Olusiji O Medaiyese, Abbas Syed, and Adrian P Lauf, “Machine Learning Framework for RF-Based Drone Detection and Identification System,” in *2021 2nd International Conference On Smart Cities, Automation & Intelligent Computing Systems (ICON-SONICS)*, Tangerang, Indonesia, Oct. 2021, pp. 58–64. doi: 10.1109/ICON-SONICS53103.2021.9617168.

[4] Qian Zhuotao, "Research on a water plant anti UAV system based on convolution neural network", Dissertation Submitted to Hangzhou Dianzi University, 2021. [Online]. Available: https://kns.cnki.net/KCMS/detail/detail.aspx?dbcode=CMFD&dbname=CMFD202201&filename=1021127966.nh&v=

[5] Ibrahim Nemer, Tarek Sheltami, Irfan Ahmad, Ansar Ul-Haque Yasar, and Mohammad A. R. Abdeen, “RF-Based UAV Detection and Identification Using Hierarchical Learning Approach,” *Sensors*, vol. 21, no. 6, p. 1947, Mar. 2021, doi: 10.3390/s21061947.

[6] Rabiye Kılıç, Nida Kumbasar, Emin Argun Oral, and Ibrahim Yucel Ozbek, “Drone classification using RF signal based spectral features,” *Eng. Sci. Technol. Int. J.*, vol. 28, p. 101028, Apr. 2022, doi: 10.1016/j.jestch.2021.06.008.

[7] Nida Kumbasar, Rabiye Kılıç, Emin Argun Oral, and Ibrahim Yucel Ozbek, “Comparison of spectrogram, persistence spectrum and percentile spectrum based image representation performances in drone detection and classification using novel hmffnet: hybrid model with feature fusion network,” *Expert Syst. Appl.*, vol. 206, p. 117654, Nov. 2022, doi: 10.1016/j.eswa.2022.117654.

[8] YANG H M, et al. UAV RF signal identification based on residual neural network[J]. Telecommunication Engineering,2023,63(1):101-106. doi: 10.20079/j.issn.1001-893x.210902003.

[9] I. W. Selesnick, “Short-Time Fourier Transform and Its Inverse”.

[10] C. Liang *et al.*, “A Kaiser Window-Based S-Transform for Time-Frequency Analysis of Power Quality Signals,” *IEEE Trans. Ind. Inform.*, vol. 18, no. 2, pp. 965–975, Feb. 2022, doi: 10.1109/TII.2021.3083240.

[11] C.-Y. Lee and T.-A. Le, “Identifying Faults of Rolling Element Based on Persistence Spectrum and Convolutional Neural Network With ResNet Structure,” *IEEE Access*, vol. 9, pp. 78241–78252, 2021, doi: 10.1109/ACCESS.2021.3083646.

[12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition.” arXiv, Dec. 10, 2015. doi: 10.48550/arXiv.1512.03385.