Identification of Aircraft Targets Based on Time-frequency Images and CNN

Ting Lu¹, Wenpeng Zhang¹, Wei Yang and Yaowen Fu⋆

¹College of Electronic Science and Technology, National University of Defence Technology, Changsha, Hunan, China

⋆Corresponding author’s e-mail: fuyaowen@sina.com

Abstract. Aimed at the micro-Doppler effect of air targets, a method is proposed based on time-frequency images and convolutional neural network (CNN) to identify helicopters, propellers and jets. In this paper, the time-frequency images are constructed by performing short-time Fourier transform on the aircraft target echo signal. Using the obtained time-frequency images as the input of the convolutional neural network, the time-frequency features of the three types of aircraft targets are extracted by utilizing the powerful image learning ability of the convolutional neural network. The results show that the CNN method has better recognition effect than the support vector machine (SVM) and back propagation (BP) network, and the CNN is more robust and generalized.

1. Introduction

Modern warfare has developed into an all-round combat mode for land, sea and air, and air combat has always been valued by all countries. Airborne aircraft targets can be divided into helicopters, propellers, and jets according to their working principles and uses. The timely and accurate identification of three types of air targets in combat is particularly important for the first opportunity to obtain air warfare.

The micro-Doppler effect was first proposed by Victor C. Chen[1-3]. The micro-Doppler of the radar target reflects the fine structure and micro-motion of the target, which is an important basis for radar target recognition. The echoes of helicopters, propellers and jets contain the fretting components from the rotating components and the body echoes of the fuselage. In response to the micro-motion of the aircraft target, Baoshuai Wang uses the EMD decomposition method to effectively separate the fretting component and the fuselage component of the jet aircraft and the helicopter target, and then uses the CLEAN method to analyse the target echoes of the three types of aircraft to achieve the three types of aircraft targets. The body component and the fretting component are completely separated[4]. Yue Jiang et al.[5]of Xi’an University of Electronic Science and Technology, based on the time-frequency images of the aircraft target fretting echo, extracted the entropy of the images, the average time-spectrum waveform entropy and the average time-frequency spectral variance into SVM to achieve the classification of three types of aircraft targets.

In recent years, The CNN successfully applied in the field of image recognition[6]. In 2012, Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton created a ‘large-scale, deep convolutional neural network’ and used it to win the 2012 ILSVRC Challenge (ImageNet Large-Scale Visual Recognition Challenge)[7]. Subsequently, many scholars proposed a variety of CNN-based network models based on different images application backgrounds, such as VGG NET[8], Google LeNet[9], Google LeNet V2[10] and a series of network models for image recognition. The CNN is used for direct processing
of 2D images. Chao Wang, Jian Wang, and Xudong Zhang of Tsinghua University proposed time-frequency images that can be identified by the CNN network[11]. Yuming Shao of the University of Science and Technology of China used the CNN network to realize the human gait micro-Doppler characteristics of the time-frequency images of human micro-motion targets[12].

The CNN mainly has the following characteristics: the original images are used as the input of the network, which avoids the complicated feature extraction process in the traditional machine learning method. The local receptive field acquires the characteristics of the input images, and the translation, scaling and rotation of the images are irrelevant, which facilitates the expansion of the images sample and can solve the network training problem of the small sample. Down sampling uses the principle of local correlation to preserve structural information while effectively reducing the amount of data processing[13]. The CNN is one of the deep learning methods, and its ‘depth’ is relative to the ‘shallow learning’ method such as SVM and maximum entropy method. Shallow learning mainly extracts the characteristics of samples through artificial experience, and the learned features are single-layer features without hierarchical structure. The convolutional neural network continuously transforms into a new feature space by transforming the original signal layer by layer, and automatically learns the deep hierarchical features. The error of manually selecting features is avoided, and the signal feature recognition rate is improved[14].

In this paper, the time-frequency images of the three types of aircraft target radar fretting echo is directly used as the input of CNN. The CNN is used to automatically complete the micro-motion feature extraction and classification of the target, which reduces the manual operation part in the feature extraction process and further improves the recognition rate of the aircraft target.

2. Time-frequency images construction of aircraft target micro-motion signal

For a Frequency modulated continuous wave (FMCW) radar, the transmitted signal is
\[ x(t) = \exp(j2\pi f_0 t) \]
where is the radar transmit signal carrier frequency. The geometric relationship of the rotating parts of the aircraft relative to the radar target is shown in Figure 1.

![Figure 1. Geometric diagram of rotating blades relative to radar. The distance from any point \( p_0(x_{r0}, y_{r0}, z_{r0}) \) on the rotating blade \( L \) to the centre \( R_0(r_0, 0, 0) \) of the rotating blade relative to the radar position \( R_0(0, 0, 0) \) is \( x \) (\( r \leq x \leq l \), \( r \) is the length from the root of the blade to the centre of the rotating blade, \( l \) is the length from the tip of the rotating blade to the centre of rotation). The initial angle between the blade and the \( x \) axis is \( \theta_0 \).

Using the method of literature[15], after the received radar echo is mixed, the fundamental frequency echo signal of a point \( p_0 \) on the rotating blade is obtained as follows:
\[ s_i(t) = \sigma_i \exp(j2\pi f_{i0} \frac{2R}{c}) = \sigma_i \exp\left[j4\pi f_{i0} \frac{R + \cos \beta x_i \cos(2\pi F_{rot} t + \theta_0)}{c}\right] \]  

(1)

Where \( R \) is the distance between the centre of the rotor's rotating blade and the phase centre of the radar, and \( \beta \) is the pitch angle. \( \sigma_i \) is a backward relative scattering coefficient, its value is set to 1.

The phase of the echo signal is \( \varphi(t) = 4\pi \frac{R + \cos \beta x_i \cos(2\pi F_{rot} t + \theta_0)}{\lambda} \), and the Doppler frequency of the rotating blade \( p_0 \) is defined as:

\[ f_d(t) = -\frac{1}{2\pi} \frac{d\varphi(t)}{dt} = \frac{4\pi F_{rot} x_i \cos \beta \sin(\theta_0 + 2\pi F_{rot} t)}{\lambda} \]  

(2)

Figure 1 shows the echo signals generated by all rotating blades of the aircraft relative to the radar:

\[ s(t) = \sum_{i=0}^{\infty} \exp(j4\pi f_{i0}(R-\nu t)) \exp(j4\pi f \frac{\cos \beta x_i \cos(2\pi F_{rot} t + \theta_0 + \frac{2\pi k}{N})}{c}) dx \]

\[ = (l-r) \exp(j \frac{4\pi R}{\lambda} \sum_{i=0}^{\infty} \sin \left( \frac{2\pi(l-r) \cos \beta \cos(2\pi F_{rot} t + \theta_0 + \frac{2\pi k}{N})}{\lambda} \right) \exp(j \left( -\frac{2\pi(l+r) \cos \beta \cos(2\pi F_{rot} t + \theta_0 + \frac{2\pi k}{N})}{\lambda} \right)) \]  

(3)

Where the wavelength of the radar signal is \( \lambda \), the number of blades is \( N \), corresponding to \( k = 0, 1, 2, \ldots, N-1 \).

In equation (3), the amplitude component of the rotating part of the aircraft is modulated by the \( \sin \) function, which determines the time domain characteristics of the echo signal. A flickering spike appears in the time domain when the blade rotates perpendicular to the radar line of sight. The time interval between time domain spikes is \( \Delta t = 2\pi/\nu N \times 2\pi F_{rot} = 1/\nu N F_{rot} \) (When an odd number is \( N \), \( n = 2 \) and when \( N \) is an even number, \( n = 1 \)). When the blade has a flashing spike in the time domain, the blade is exposed to the radar with the highest intensity, and each scattering point on the blade produces a micro-Doppler frequency, a flashing band of \( \frac{4\pi F_{rot} R}{\lambda} \cos \beta \) \( - \frac{4\pi F_{rot} l}{\lambda} \cos \beta \) is expressed in the frequency domain. Corresponding frequency domain unilateral spectrum number is \( N_1 = \frac{8\pi(l-r) \cos \beta}{N\lambda} \), line spacing is \( \Delta f = N \times F_{rot} \).

Aircraft radar target echo signals are nonlinear and non-stationary, signal time-frequency domain information obtained by short-time Fourier transform (STFT), its STFT is defined as:

\[ STFT(t, f) = \int_{-\infty}^{+\infty} s(\tau)h(\tau-t)e^{-j2\pi ft} d\tau = \left\langle s(\tau), h(\tau-t)e^{-j2\pi ft} \right\rangle \]  

(4)

Where \( s(t) \) is the original signal, \( h(t) \) is the window function. STFT is performed on the echo signals of helicopter, propeller and jet to obtain time-frequency images.
The simulation parameters of the three types of aircraft targets correspond to the numbers A in Table 1, Table 2, and Table 3, respectively. The radar carrier frequency is $f_0 = 5GHz$ and the beam dwell time is 400 ms. When $\beta = \frac{\pi}{6}$, signal noise ratio $SNR = 30$ dB, the time-frequency images of helicopters, propellers and jets are shown in Figure 2 (a)(b)(c). Three aircraft target time-frequency images. For helicopters, the strobe time interval is $\Delta t = 25.4$ ms, the maximum Doppler frequency is $f_{d_{\text{max}}} = 6366.2$ Hz, the number of single-sided spectrums in the Doppler domain is $N_1 = 646$, and the line spacing is $\Delta f = 123.8$ Hz. For propeller aircraft, strobe time interval $\Delta t = 12.5$ ms, maximum Doppler frequency $f_{d_{\text{max}}} = 6166.9$ Hz, minimum Doppler frequency $f_{d_{\text{min}}} = 2466.8$ Hz, Doppler domain unilateral spectrum number $N_1 = 92$, line spacing $\Delta f = 502.7$ Hz. For jet aircraft, the strobe interval is $\Delta t = 0.444$ ms, the maximum Doppler frequency is $f_{d_{\text{max}}} = 1.3516 \times 10^4$ Hz, the number of single-sided spectrum in the Doppler domain $N_1 = 6$, and the line spacing $\Delta f = 2229.3$ Hz. The above theoretical calculation results are basically consistent with the simulation results in the figure.

3. Time-frequency images convolutional neural network
The CNN is a multi-layer neural network structure designed as a deep feedforward neural network specifically for extracting two-dimensional signal features. The TFI-CNN network model is proposed to learn and identify the time-frequency images[16], which is suitable for the feature learning and recognition of the aircraft target time-frequency images. The CNN model is designed for the aircraft target time-frequency images.
Figure 4. This is a CNN network structure parameter flow chart description. The designed CNN network model consists of three convolutional layers, three pooling layers and one fully connected layer. The number of convolution kernels per layer of convolutional layers is 20, 40, and 60, respectively. Each convolution kernel has a size of $3 \times 3$ and a stride of 1. The pooling layer uses the maximum pooling, the size of the kernel is $2 \times 2$, and the stride is 2 steps. The activation functions used are all ReLU, and the last connection is the softmax layer.

4. Simulation experiment results and analysis

4.1 Simulation experiment parameters

Using the simulation parameters of helicopters, propellers and jets in Table 1, Table 2 and Table 3, which obtains the time-frequency images of the target rotating part of the aircraft. Set the radar working carrier frequency $f_0$, beam dwell time 400. By changing the aircraft's pitch angle $\beta$, we obtained 50 time-frequency images samples for each of the three types of aircraft, for a total of 900 images sample data. Randomly select 180 of them as the training set, and the rest as the test set. The original images of the generated time-frequency images are input to the CNN network.

In this experiment, under the Windows system, the Tensorflow deep learning toolkit is used to identify the time-frequency images based on the GPU. The model parameters are continuously selected and optimized, and the optimized parameter settings are shown in Table 4.

| Model | A  | B  | C  | D  | E  | E  |
|-------|----|----|----|----|----|----|
| Rotor speed (rpm) | 394 | 242 | 217 | 395 | 132 | 394 |
| $r$ /m | 0 | 0 | 0 | 0 | 0 | 0 |
| $l$ /m | 5.345 | 8.6 | 7.8 | 5.335 | 16 | 5.64 |
| Number of blades(N) | 3 | 5 | 4 | 4 | 8 | 32 |

| Model | A | B | C | D | E | F |
|-------|---|---|---|---|---|---|
| Rotor speed (rpm) | 1200 | 3400 | 2340 | 1245 | 3400 | 7010 |
| $r$ /m | 0.68 | 0.27 | 0.23 | 0.79 | 0.23 | 0.12 |
Table 3. Jet simulation parameters

| Model  | A     | B     | C     | D     | E     |
|--------|-------|-------|-------|-------|-------|
| Rotor speed (rpm) | 3520  | 3000  | 4000  | 8615  | 5000  |
| \( r / \text{m} \) | 0.38  | 0.3   | 0.24  | 0.18  | 0.2   |
| \( l / \text{m} \) | 0.1   | 1     | 0.8   | 0.51  | 0.6   |
| Number of blades(N) | 38    | 30    | 42    | 27    | 33    |

Table 4. CNN experiment parameter setting

| Parameter          | Value |
|--------------------|-------|
| Number of iterations \( (i) \) | 100   |
| Learning rate \( \varepsilon \) | 0.001 |
| Mini-batch(D)      | 50    |

4.2 Simulation experiment results

4.2.1 Noise-free multi-sample recognition results

The number of training samples is 180, and the number of test samples is 720, regardless of noise. The variation rate of CNN's recognition rate of the time-frequency images generated by the micro-motion of three aircraft target rotating parts with the number of iterations is shown in Figure 5.

![Figure 5. Recognition rate curve.](image)

![Figure 6. Recognition rate of different samples.](image)

It can be seen from Figure 5 that the time-frequency images features of the three aircraft targets are more obvious. As the number of iterations increases to 2, the recognition rate reaches 100% under the training set. Compared with the literature[5], the time-frequency images are used to extract the time-frequency images entropy and the SVM classifier is used. In the ideal case, the recognition rate is 95.53%. In this paper, the frequency domain recognition effect is obviously improved in the three aircraft targets. After 50 repeated experiments, the number of samples selected in each experiment was different. After multiple iterations, the recognition rate was still 100%, and the performance was very stable.

4.2.2 Effect of sample size on recognition result

In the actual battlefield situation, the micro-motion characteristics of the aircraft target are difficult to capture, and the number of samples collected is relatively small. We use small samples for training
and consider the generalization ability of the CNN model. The number of training samples of the three aircraft targets was changed for experiment. Two samples of each aircraft were randomly selected as the training set, and the remaining samples were used as tests. Under different SNR conditions, the corresponding recognition results were obtained as shown in Figure 6.

In Figure 6, in the case of small samples, the CNN model can obtain a better recognition rate when the SNR is high. When the SNR is low, the learning ability of the CNN is reduced, and the recognition effect is significantly reduced. In the case of multiple samples, the CNN is getting a higher recognition rate, and the influence of noise on its recognition rate is small.

4.2.3 Influence of SNR on recognition results

![Figure 7. CNN recognition rate under different SNR.](image)

![Figure 8. Recognition rate of three classifiers with different SNR.](image)

For the actual situation, the model is required to be robust to noise. A certain amount of Gaussian white noise is added to the original signals of the three aircraft targets. As can be seen in Figure 7, the CNN network is robust to noise. Even in the case of -5dB Gaussian white noise, the recognition rate can still reach 99.12% after 19 iterations. In Figure 8, the recognition rates of SVM, BP network and the CNN in the case of different signal to noise ratios are compared. SVM and BP use the characteristics of the time-frequency images entropy of three aircraft targets as training test samples. As the SNR increases, the CNN network recognition effect is significantly higher than the SVM and BP network recognition rates. In the case of low signal to noise ratio, it also shows a good recognition effect.

5. Conclusion

In this paper, the influence of Doppler generated by the movement of the target body of the aircraft is compensated, and the micro-motion of the three rotating parts of the aircraft is considered separately by the fretting extraction algorithm to remove the fuselage echo, and their time-frequency images are constructed. The CNN is used to classify the features of the time-frequency images directly, and a good recognition rate is obtained. Compared with the SVM and BP, the CNN performs better. In addition, considering the complexity of the actual situation, through the test of the addition of noise and small sample CNN model, it is proved that the CNN model can improve the SNR to train a better model to achieve aircraft target recognition under small sample conditions, which has a certain degree of robustness and generalization capabilities.

References

[1] Chen V C, Li F, Ho S S, et al. (2006) Micro-Doppler effect in radar: phenomenon, model, and simulation study. J. IEEE Transactions on Aerospace and Electronic Systems, 42(1):2-21.

[2] Chen V C. (2003) Micro-Doppler effect of micro motion dynamics: a review. Proceedings of SPIE - The International Society for Optical Engineering, 5102:240-249.

[3] Chen V C, Li F, Ho S S, et al.(2003) Analysis of micro-Doppler signatures. C. IEE Proc. Radar Sonar Navig, 150 (4): 271-276.
[4] Baoshuai Wang. (2015) Airborne target classification based on micro-Doppler effect. D.
[5] Yue Jiang. (2014) Research on time-frequency domain feature extraction method of aircraft target based on micro-Doppler. D. Xidian University.
[6] Lin Z., Ji K., Kang M., et al. (2017) Deep Convolutional Highway Unit Network for SAR Target Classification With Limited Labeled Training Data. J. IEEE Geoscience and Remote Sensing Letters: 1-5.
[7] Krizhevsky A, Sutskever I, Hinton G E. (2012) ImageNet classification with deep convolutional neural networks. C.// International Conference on Neural Information Processing Systems. Curran Associates Inc: 1097-1105.
[8] Simonyan K, Zisserman A. (2014) Very deep convolutional networks for large-scale image recognition. J. arXiv preprint arXiv: 1409.1556.
[9] Szegedy C, Liu W, Jia Y, et al. (2014) Going deeper with convolutions. J.
[10] Szegedy C, Vanhoucke V, Ioffe S, et al. (2016) Rethinking the Inception Architecture for Computer Vision. C.// Computer Vision and Pattern Recognition. IEEE: 2818-2826.
[11] Wang C, Wang J, Zhang X. (2017) Automatic radar waveform recognition based on time-frequency analysis and convolutional neural network. C.// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE: 2437-2441.
[12] Yuming Shao. (2018) Human micro-motion feature recognition based on deep learning method. D. University of Science and Technology of China.
[13] Baocai Yin, Wentong Wang, Lichun Wang. (2015) Review of Deep Learning Research. J. Journal of Beijing University of Technology, (1): 48-59.
[14] Lecun Y, Bengio Y, and Hinton G. (2015) Deep learning. J. Nature, 521(7553): 436-444.
[15] Guangfeng Chen. (2014) Radar target micro-Doppler feature analysis and its application. D. Xidian University: 78-79.
[16] Wang C, Wang J, Zhang X. (2017) Automatic radar waveform recognition based on time-frequency analysis and convolutional neural network. C.// IEEE International Conference on Acoustics, Speech and Signal Processing. IEEE: 2437-2441.