The Journal of Engineering

IET International Radar Conference (IRC 2018)

eISSN 2051-3305
Received on 19th February 2019
Accepted on 02nd May 2019
E-First on 17th July 2019
doi: 10.1049/joe.2019.0179
www.ietdl.org

Motion classification for radar moving target via STFT and convolution neural network

Xiaoqian Mou[1], Xiaolong Chen[1] ✉, Ningyuan Su[1], Jian Guan[1]
[1]*Radar Target Detection Research Group, Naval Aviation University, Erma Road 188, Yantai, People's Republic of China*
✉ *E-mail: cxlcxl1209@163.com*

Abstract: Radar moving target detection (MTD) technology is a key technology in the field of radar signal processing. The MTD method of radar has a better performance when applied to uniform motion detection although it has limited performance in other aspects, and it is also difficult to distinguish the type of the moving target. This article presents a new method for detecting and classifying moving targets based on convolutional neural network, which uses the convolutional neural network to learn the motion characteristics of the moving target echo so that realises the detection and classification of moving targets. At first, the authors model the moving target echo. Then the short-time Fourier transform (STFT) is used to perform time–frequency analysis. The obtained time–frequency graph is used as the input of the convolutional neural network which can automatically learns moving target features through training. Finally, the authors use the trained convolutional neural network model to detect and classify moving targets. The simulation verifies that the method has a great improvement in the detection accuracy rate compared with the traditional MTD.

1 Introduction

Moving target detection (MTD) is the key point in the detection and research of radar, which play an important role on searching for the target, monitoring target and tracking target. MTD research also has been an important branch of radar target detection. The detection of moving target is influenced by target mobility, clutter and noise disturbance, which is difficult in the actual detection under complicated background. The traditional MTD method is MTD. For the first, it extracts Doppler information, then filter out noise clutter by Doppler filter group for motion target information, achieving the function of target detection. There are still some shortcomings on MTD's detection performance. MTD applies only better to uniform motion target. For variable speed moving targets, MTD has the problem of spectral divergence and difficulty in energy accumulation. It is also difficult for multi-component signal processing of different types of motion.

In recent years, with the rapid development of artificial intelligence, deep learning has also been increasingly researched and applied in the field of target recognition. The convolutional neural network has great advantages in image recognition and target detection [1]. It can automatically extract image features through convolution kernel to achieve a good target recognition function and achieve a higher detection success rate. In 1989, Professor Yann LeCun of the University of Toronto in Canada proposed a convolutional neural network together with his colleagues. In 2012, Hinton improved the training method of the network, after that the convolutional neural network achieved a huge breakthrough in image target detection in the field of target detection. In 2016, Tian *et al.* proposed a target recognition research of synthetic aperture radar (SAR) image based on convolutional neural network [2]. In 2017, Wang Siyu proposed a plane target detection algorithm of SAR image with high resolution based on convolutional neural network, which is a new model of aircraft target detection [3]. However, the above studies are based on SAR radar imaging research and have certain limitations on the radar system.

In this paper, we propose a MTD technique based on convolutional neural network time–frequency processing. At first, the target echo signal is extracted the target Doppler shift information after demodulation and pulse compression. Then we transformed it into a time–frequency graph through the short-time Fourier transform, and input it to the convolutional neural network, using its powerful image recognition function to detect moving target. Through this method, we can identify uniform motion, uniform speed, micro-moving target, and make up for the traditional detection method MTD's limitations.

2 Echo demodulation and construction of time–frequency spectrum

2.1 Echo demodulation and matched filtering

Assume that the radar transmits a linear frequency modulation signal

$$s(t) = \mathrm{rect}\left(\frac{t}{T}\right)\mathrm{e}^{j2\pi\left(f_c t + \frac{1}{2}ut^2\right)}, \tag{1}$$

where f_c is the carrier frequency, $\mathrm{rect}(t/T)$ is a rectangular signal, T is the width of pulse, $u = B/T$ is modulation frequency, and B is bandwidth

$$\mathrm{rect}\left(\frac{t}{T}\right) = \begin{cases} 1, & \frac{t}{T} \le 1 \\ 0, & \text{other} \end{cases}. \tag{2}$$

For the received target echo, we extract the Doppler shift information of uniform motion, uniform variable motion, and micro motion target echo signal through demodulation and pulse compression [4]

$$f_1 = \frac{2}{\lambda}(v_0 + a_s t_m), \tag{3}$$

$$f_2 = \frac{2v_0}{\lambda}, \tag{4}$$

$$f_3 = \frac{2Aw\cos(wt_m + \varphi)}{\lambda}, \tag{5}$$

where f_1, f_2, and f_3 is the frequency information of target echo extracted by demodulation and pulse compression, v_0 is the initial speed of the target, a_s is the acceleration of the target, t_m is pulse-to-pulse slow time during coherent processing intervals, A, w, and

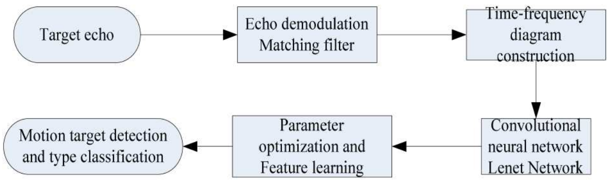

Fig. 1 *Block diagram of the model*

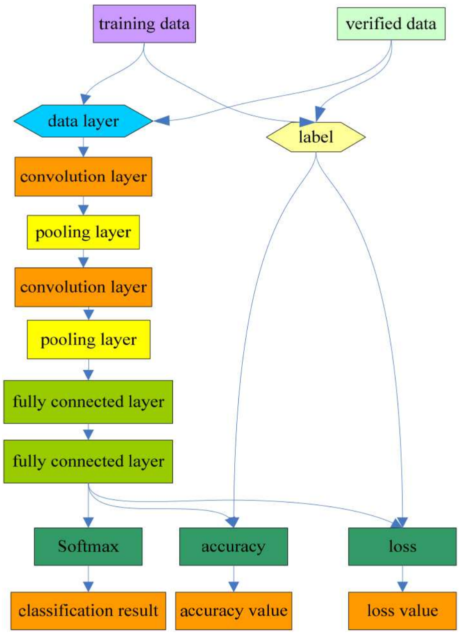

Fig. 2 *CNN structure*

φ are sinusoidal amplitude, angular frequency and phase of the jog target, and λ is radar wavelength.

2.2 Doppler shift information to time–frequency map conversion

It can be seen from the moving target echo model that the echo has time-varying property and non-stationary, so it should be analysed by the method of time–frequency analysis. The common method is short-time Fourier transform (STFT) and Wigner-Ville distribution (WVD). WVD has a very high resolution, but it will be difficult to avoid cross-terms in the multi-signal analysis, so we use STFT for time–frequency analysis [5]. We use the short-time Fourier transform to convert the extracted Doppler shift information into two-dimensional information, which is converted into a time–frequency graph by MATLAB

$$\mathrm{STFT}(t, f) = \int_{-\infty}^{+\infty} [S_{\mathrm{m}}(t)g^*(u-t)]e^{-j2\pi f u}\mathrm{d}u, \tag{6}$$

where $S_{\mathrm{m}}(t)$ is motion target echo after demodulation and pulse compression and $g(t)$ is Hanning window.

3 Establishment of convolutional neural network (CNN) and moving target detection method

3.1 Establishment of convolutional neural network model

The convolutional neural network (CNN) consists of convolutional layer, pooled layer and fully connected layer. The convolutional layer is the core of the convolutional neural network. The features of the image are extracted through a convolution kernel (filter). The pooling layer is actually downsampling, reducing the amount of computation and data dimensions and the full connection layer can convert the output of the previous layer into a vector.

The advantage of convolutional neural network training is that it can use the error of the output value and the true value to carry out the back propagation so as to adjust the weight coefficient to optimise the convolutional neural network and further improve the accuracy of forward propagation. The loss function is used to represent the error between the output value and the true value. The loss function is established as follows:

$$E^N = \frac{1}{2}\sum_{n=1}^{N}\sum_{q=1}^{c}(t_q^n - X_q^n)^2, \tag{7}$$

where N is the number of samples, c is the number of sample categories, t_q^n is the k dimension of the n Sample real value, X_q^n is the k dimension of the n Sample output value, w is multiplicative coefficient, and b is offset factor.

Our objective function is to find the minimum value of the loss function so that the convolutional neural network model is optimised

$$\min E^N. \tag{8}$$

In this paper, we selected the caffe architecture for the study of the model. At the same time, this is a four-category problem of uniform speed, uniform acceleration, uniform deceleration, and micro-motion targets, so we chose the Lenet network as the basic model network. The network can do a good job of solving 10 classification problems of handwritten numbers.

3.2 CNN-based moving target detection and classification

The flow chart of this model is as follows. As shown in Fig. 1, at the beginning we perform echo demodulation and matched filtering on the moving target echo of the simulation to complete the extraction of moving target Doppler information [6]; then we use STFT to analyse the moving target echo information by time–frequency, which is to construct the time–frequency diagram. Finally, the time–frequency diagram is inputted into the convolutional neural network to get the characteristic information of various moving targets. The parameters of the network are optimised through iteration so that CNN MTD and classification model is established [7].

The structure of the convolutional neural network in this paper is shown in Fig. 2. We set the input image data as 128 pixels, and finally realise the four-classification function for uniform speed, uniform acceleration, uniform deceleration, and micro-motion target. CNN network structure diagram and network parameter setting table and parameter scale table are shown in Tables 1 and 2.

Table 1 Network parameter table

Parameter	Value
number of iterations	30
optimal parameter solving algorithm	SGD
initial learning rate	0.01
learning strategy	step
step size	109
gamma	0.1

Table 2 Network parameter scale

Primary layer name	Network parameter scale
convolution 1	1520
convolution 2	25,050
fully connected layer 1	21,025,500
fully connected layer 2	2004

4 Simulation analysis and test

4.1 Construction of a simulated image dataset

This paper builds two data sets. The first data set contains 2400 images with 128 pixels, which includes 600 pieces of time–frequency maps for uniform speed, uniform deceleration, uniform

J. Eng., 2019, Vol. 2019 Iss. 19, pp. 6287-6290

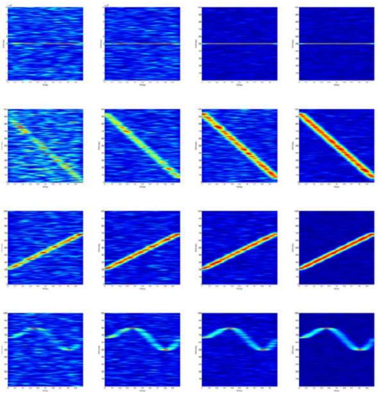

Fig. 3 *Data set simulation example*

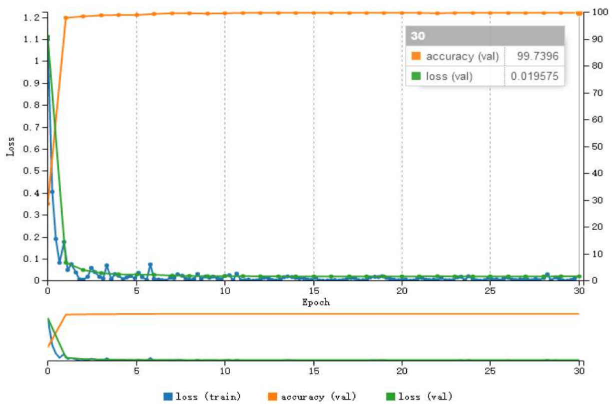

Fig. 4 *Training model results*

acceleration, and jog target. The second data set contains 240 images with 128 pixels, which includes 60 pieces of time–frequency maps for uniform speed, uniform deceleration, uniform acceleration, and jog target. The Doppler shift value of the simulation signal is between 100 Hz and 10 kHz. The echo signal has Gaussian noise and the signal-to-noise ratio (SNR) is −5 to 9 dB. The type of time–frequency diagram in the data set is shown in Fig. 3.

4.2 Convolutional neural network moving target recognition results

We created a environment based on VS2013, CUDA7.5, cudnn 5.1, caffe and digits, while digits is a graphical interface operation tool that supports caffe. After the first data set was input into the convolutional neural network model, the results are shown in Fig. 4.

We set 75% of time–frequency graph for training and 25% for verification. As shown in Fig. 4, the brown curve is the recognition accuracy curve. The accuracy rate at the 30th iteration is 99.796%. The green curve is the curve of the verification loss function value, and the loss value is as low as 0.00001. (The smaller the loss value, the better the objective function of the solution.)

Fig. 5 shows the learning rate curve. The smaller the learning rate value is, the smaller the step size is and the higher the precision is. When the number of iterations is after the 20th time, the learning rate is close to zero, indicating that the accuracy is very high.

4.3 Single image test

When the number of training samples are 2400, we select uniform speed, uniform acceleration, uniform deceleration, and micro-motion target time–frequency graph with a SNR of −1 and the size of 128 × 128 to test. The results are shown in Table 3.

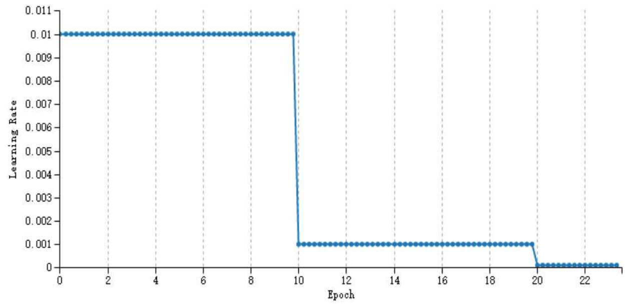

Fig. 5 *Training model learning rate*

Table 3 Detection and classification results of four different types of moving target time–frequency maps

Recognition rate	Uniform speed, %	Uniform acceleration, %	Uniform deceleration, %	Micro-motion, %
uniform speed	91.36	7.84	0.8	0
uniform acceleration	0	99.92	0.08	0
uniform deceleration	0	0.02	99.98	0
micro-motion	0	0	0	100

4.4 Test with mass 128 × 128 images

In order to study the impact of SNR on target detection performance, we used the target time–frequency graph with SNR of −8 to 6 dB for each type of moving target to test. The results are shown in Tables 4 and 5.

From the results of the above two tables, it can be seen that the uniform moving target is relatively difficult to detect, and its detection probability increases with the increase of the SNR. The detection probability of uniform acceleration, uniform deceleration, and non-uniform motion targets is quite high and accurate. With the increase of the number of samples, the detection probability of the uniform motion of the network at the same SNR is also improved. The detection performance of the network for the micro-motion target is good, and the SNR is the least affected. Therefore, the deep convolutional neural networks is more suitable for handling complex motion type targets.

4.5 Comparison with traditional methods

In order to test the superiority of convolutional neural networks, we use the traditional MTD method MTD to conduct a comparative simulation study.

From Table 6, it can be seen that the MTD technology based on convolutional neural network time–frequency processing is effective, and the detection performance and detection probability are superior to the traditional MTD method.

5 Conclusions

In this paper, we use the combination of STFT and convolutional neural network to convert the target information into time–frequency map information and put it into the convolutional neural network, which is a good realisation of radar uniform, uniform acceleration, micro-MTD. We propose a new MTD technology based on the convolutional neural network time–frequency map processing, making up for the defects of the traditional detection method MTD and improving the detection probability, and apply deep learning to radar target detection so that it broadens ideas of the radar target recognition.

6 Acknowledgments

This work was supported in part by the National Natural Science Foundation of China (61501487, U1633122, 61471382, and 61531020), National Defense Science Foundation (2102024),

Table 4 Detection and classification result table when the total number of training samples is 240 (60 in each category)

Type	SNR, dB					
	−10	−5	−2	0	2	6
uniform speed, %	0	0.0	0.94	50.9	84.9	99.38
uniform acceleration, %	0.4	6.87	88.09	100	100	100
uniform deceleration, %	0.38	6.24	86.75	100	100	100
micro-motion, %	0.34	100	100	100	100	100

Table 5 Test results and classification results when the total number of training samples is 2400 (600 in each category)

Type	SNR, dB					
	−10	−5	−2	0	2	4
uniform speed, %	0	2.04	83.57	91.95	98.09	99.73
uniform acceleration, %	93.09	99.6	99.99	100	100	100
uniform deceleration, %	90.51	99.4	100	100	100	100
micro-motion, %	92.84	100	100	100	100	100

Table 6 Comparison of test performance

Moving target detection method	Recognition rate, %	Detection performance
convolutional neural network time–frequency map	99.80	constant speed, uniform speed, micro-motion target
MTD	88.34	uniform target

Young Elite Scientist Sponsorship Program of CAST and Special Funds of Taishan Scholars of Shandong.

7 References

[1] Krizhevsky, A., Sutskever, I., Hinton, G.E.: 'Imagenet classification with deep convolutional neural networks', *NIPS Curran Associates Inc.*, 2012, **25**, (2), p. 2012

[2] Tian, Z.Z., Zhan, R.H., Hu, J.M.*, et al.*: 'Research on SAR image recognition based on convolutional neural network', *J. Radar*, 2016, **5**, (3), pp. 321–324

[3] Wang, S.Y., Gao, X., Zheng, X.W.*, et al.*: 'High resolution SAR image aircraft target detection method based on convolutional neural network', *J. Radar*, 2017, **6**, (2), pp. 196–201

[4] Chen, X., Guan, J., Huang, Y.*, et al.*: 'Radon-linear canonical ambiguity function-based detection and estimation method for marine target with micromotion', *IEEE Trans. Geosci. Remote Sens.*, 2015, **53**, (4), pp. 2225–2240

[5] Allen, J.B., Rabiner, L.R.: 'A unified approach to short time Fourier analysis and synthesis', *Proc. IEEE*, 1977, **65**, pp. 1558–1564

[6] Chen, X., Guan, J., Bao, Z.*, et al.*: 'Detection and extraction of target with micro-motion in spiky sea clutter via short-time fractional Fourier transform', *IEEE Trans. Geosci. Remote Sens.*, 2014, **52**, (2), pp. 1002–1018

[7] Srivastava, N., Hinton, G., Krizhevsky, A.*, et al.*: 'Dropout: a simple way to prevent neural networks from overfitting', *J. Mach. Learn. Res.*, 2014, **15**, (1), pp. 1929–1958