

# Radar-ID: human identification based on radar micro-Doppler signatures using deep convolutional neural networks

ISSN 1751-8784

Received on 11th November 2017

Revised 8th March 2018

Accepted on 17th March 2018

E-First on 13th April 2018

doi: 10.1049/iet-rsn.2017.0511

www.ietdl.org

Peibei Cao<sup>1,2</sup>, Weijie Xia<sup>1,2</sup>✉, Ming Ye<sup>1,2</sup>, Jutong Zhang<sup>1,2</sup>, Jianjiang Zhou<sup>1,2</sup>

<sup>1</sup>Key Laboratory of Radar Imaging and Microwave Photonics, Ministry of Education, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, People's Republic of China

<sup>2</sup>College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, People's Republic of China

✉ E-mail: nuaaxwj@nuaa.edu.cn

**Abstract:** Human identification is crucial in various applications, including terrorist attack preventing, criminal seeking, defence and so on. Traditional human identification methods are usually based on vision, biological features, radio-frequency identification cards and so on. In this study, the authors propose an identification method based on radar micro-Doppler signatures using deep convolutional neural networks (DCNNs) for the first time, which can identify human in non-contact, remote and no lighting status. They employ a K-band Doppler radar to acquire the raw signals due to its stationary clutter rejection and movement detection ability as well as its short wavelength which can generate larger Doppler shift. Then short-time Fourier transform is applied to the raw signals to characterise micro-Doppler signatures. They adopt the DCNNs to deal with the spectrograms for human identification problem. The DCNNs can learn the necessary features and classification conditions from raw micro-Doppler spectrograms without employing any explicit features. While the traditional supervised learning techniques relying on the extracted features require domain knowledge of each problem. It is shown that this method can achieve average accuracy ~97.1% for 4 people, 90.9% for 6 people, 89.1% for 8 people, 85.6% for 10 people, 77.4% for 12 people, 72.6% for 16 people and 68.9% for 20 people.

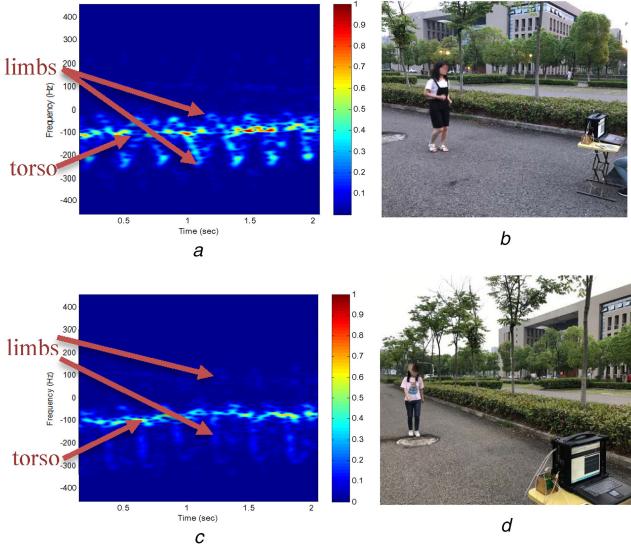
## 1 Introduction

As the increasing concerns of security and surveillance, human identification is getting more and more attention [1, 2]. In prior works, a variety of methods have been investigated for personnel recognition [3–8]. These methods usually based on biological features like fingerprints or rely on sensor technologies such as vision and sound. Since individual's activity is unique, some wearable sensors are also used to identify people. Although these methods have achieved high accuracy, they still have their own limitations. Wearable sensors [9] require to be worn in a strict way to ensure correct operations, whereas most people are uncomfortable with a sensor around them all the time. Methods based on video or camera [5, 7] usually need enough light and line-of-sight conditions. Besides, these methods invade people's privacy to a certain extent which is similar to sound-based [3, 6] methods. Additionally, a plenty of researches focus on biological features including fingerprints [10], sclera and so on have been done. Though the accuracy of these approaches is higher than other methods, collecting data is very tough and not friendly with people, which limits their applications. Meanwhile, it has been verified that biological features can be forged. Recently, some works using WIFI signal to sense human body have been done. Zhang and Bo [11] utilised channel state information to identify human. However, the wavelength of WIFI is longer than K-band radar which results in the bad performance of WIFI system in the detection of human motion. Meanwhile, the directivity of WIFI system is poor which means the system can be disturbed easily.

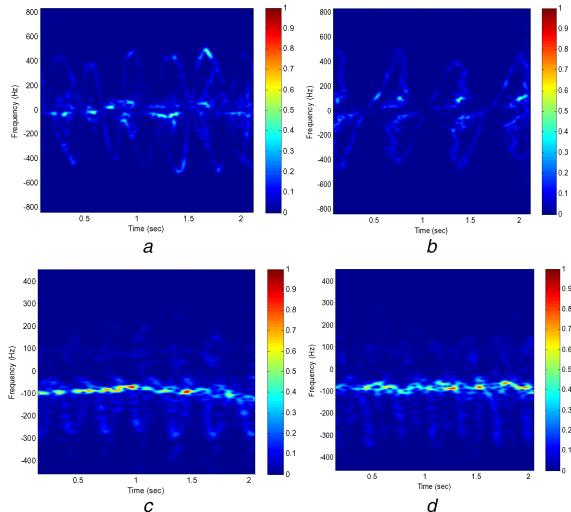
In this paper, we show for the first time that the deep convolutional neural networks (DCNNs) can be an alternative method to deal with human identification problem based on micro-Doppler signature. Micro-Doppler features produced from limb motions of individuals are unique [12] which play a pivotal role in the process. Furthermore, micro-Doppler signatures have led to various applications for recognition of human, activities and species. In [13], a comprehensive review of micro-Doppler signatures relating to different kinds of targets was given. It also

reviewed the importance and applications of micro-Doppler signatures. Seyfioğlu *et al.* [14] utilised a three-layer, deep convolutional autoencoder to distinguish seven gaits based on micro-Doppler signatures collected by a 4 GHz continuous wave (CW) radar. In [15], a Naïve Bayesian classifier and a shape-similarity-spectrum classifier were employed to recognise human based on a set of shape spectrum features extracted from the cadence velocity diagram of the human micro-Doppler signature collected by an X-band radar. Fioranelli *et al.* [16] utilised a multistatic radar system to perform recognition of people walking and three different sets of features are tested. In [17], a comprehensive survey of micro-Doppler features and their dependence upon system parameters as well as operational conditions was given. Doppler radar has been suggested as a suitable tool to collect data for its resistance to bad light and weather conditions. Specifically, compared with the signals of optical system and video surveillance system, radar signals are able to penetrate barriers such as walls and clothing with lower propagation loss. Besides, a Doppler radar is easy to build and cost effective, which makes it easy to be used widely. The same activity performed by individuals will create different perturbations in the spectrograms [18]. We can observe the difference in Fig. 1 which shows the spectrograms of two people running in the same place.

We expect to yield good results in human identification problem by exploiting the DCNNs. The DCNNs is a kind of mathematical model imitating biological neural networks, which is similar to the brain neural synaptic connections. The input of the DCNNs is the original two-dimensional images while the output is the classification results. It observably outperforms traditional supervised learning techniques that rely on the extracted features and require domain knowledge of each problem, in several applications like pattern recognition, image processing and speech recognition [19–21]. The reason for such success is that the DCNNs is able to learn the necessary features and classification conditions from raw micro-Doppler spectrograms directly [22]. Also, its significant characteristics including non-linear, high parallelism, robustness and dealing with imprecise information are



**Fig. 1** Sample spectrograms of different people  
(a) Spectrogram of (b), (c) Spectrogram of (d)



**Fig. 2** Sample spectrograms of different activities and different people  
(a) Spectrogram of a person boxing, (b) Spectrogram of a person swinging arms, (c), (d) Spectrograms of two different people running

accounted. In [22], the authors proposed the use of the DCNNs to recognise different activities. However, it is easier to classify human activities than identify individuals, because the spectrograms of different human activities are significantly different while spectrograms of different people with the same activity are very similar, as shown in Fig. 2. Yang and Lu [12] employed time-frequency transforms and pattern recognition techniques to classify individuals based on their movements with a 10 GHz radar. So far as we know, the effectiveness of identifying individuals based on radar micro-Doppler signatures using DCNNs has not been investigated. We will present our experimental results and brief introduction on the DCNNs.

The remaining paper is organised as follows. Section 2 illustrates experiments setup and data processing. Brief introduction on deep learning and the DCNNs will be described in Section 3. In Section 4, the experiment results and discussion will be presented. Section 5 concludes the paper.

## 2 Experiment setup and data processing

### 2.1 Experiment setup

The experiments were performed in the parking lot of the campus, which is shown in Fig. 3a. While Fig. 3b shows the equipment and system deployment of our experiments. The system which includes IVS-179 radar, M2i.4912 eight-channel parallel data acquisition



**Fig. 3** Data collection setup and experiment scenario

card and ACME industrial personal portable computer, is employed to collect experimental micro-Doppler signatures of people running. The IVS-179 Doppler radar works at 24 GHz in the CW mode without modulation while the ACME data recording industrial samples at 2 kHz.

Twenty-four persons participated in the data collection. There were 12 males with the maximum height 186 cm and minimum height 168 cm, 12 females with the maximum height 172 cm, compared to the minimum height 157 cm. The average weight of the males and the females were about 75 and 48 kg, respectively, and they were all at the age from 24 to 28 years. During the process, there was no disturbance created by other people. Each person walked along the line-of-sight path of the radar for 50 times and at each time a person walked for 10 s. Participants in the experiments were required to start from approximately 10 m in front of the radar.

### 2.2 Data processing

According to the Doppler effect, a moving target relative to the wave source will cause a change in frequency or wavelength of a wave. Meanwhile, if the moving target has rotating or vibrating parts, the additional frequency components in addition to the main Doppler shift will be observed, which are called the micro-Doppler effect. As human have limb motions when they are moving, micro-Doppler will be generated in radar signatures which can be clearly observed in the joint time–frequency space [23]. Therefore, short-time Fourier transform (STFT) is exploited to characterise micro-Doppler signatures.

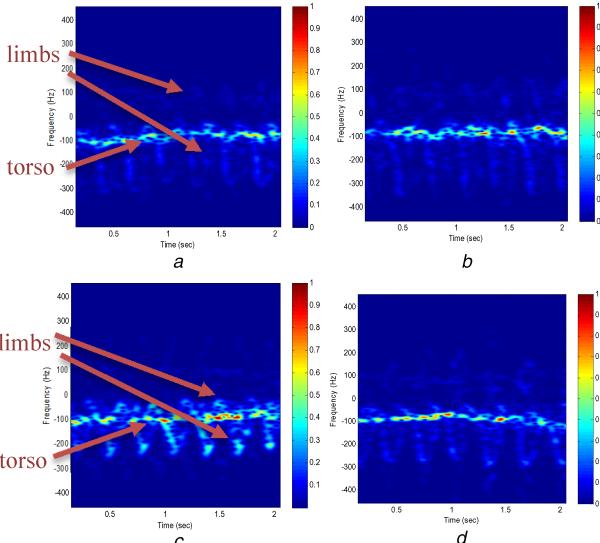
If window function is  $g(t)$  which will slide along the time line, the time-domain Doppler signal is designated as  $x(t)$ ; its STFT [24] can be expressed as

$$F_{\text{STFT}_x}(t, f) = \int_{-\infty}^{+\infty} x(u)g^*(u-t)e^{-j2\pi fu}du \quad (1)$$

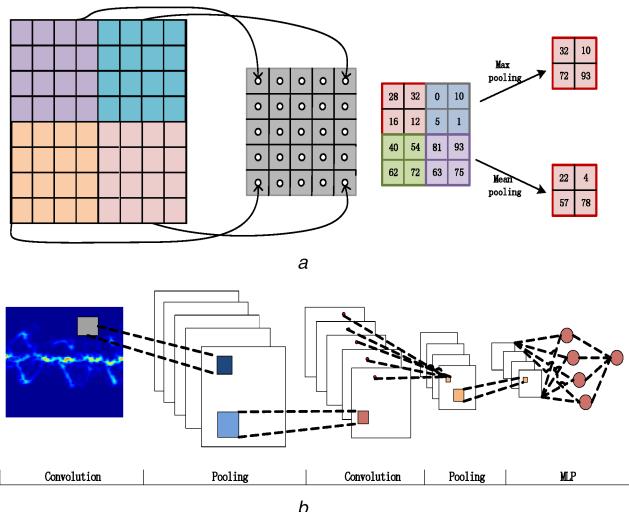
In our work, we choose Gaussian window, then  $F_{\text{STFT}_x}(t, f)$  can be expressed as

$$F_{\text{STFT}_x}(t, f) = \int_{-\infty}^{+\infty} x(u)e^{-((t-u)/2\sigma^2)}e^{-j2\pi fu}du \quad (2)$$

Proper time-window size and sliding step is vital in capturing particular features of targets in Doppler domain. After repeated practice, we choose 0.132 s as the time-window size and the sliding step size is 1/2000 s, which is appropriate to recognise the micro-Doppler characteristics in the frequency domain. Then to expand the dataset, data for one test is divided into five parts by average. The resulting spectrograms of four subjects are shown in Fig. 4. In order to get sufficient data, we employ translation, rotation, zoom, mirroring and cropping, the basic image transformation methods, to the spectrograms. These methods are often used to augment the dataset, as the basic nature of the images would not be altered by these transformations, which means the classification results would not be influenced. After these operations the number of spectrograms for each person is 4000. Furthermore, we are expanding the dataset actively for future researches and applications.



**Fig. 4** Sample spectrograms of different runners



**Fig. 5** DCNNs architecture and internal operations

(a) Schematic of the convolution filter and the pooling operation, (b) Simple DCNNs architecture

As shown in Fig. 4, the four spectrograms depending on the people are different from each other. In the spectrograms the strongest returns come from the torso, while the periodic waveforms surrounding the torso echo come from limb movements. We can clearly observe from the four spectrograms in Fig. 4 that the energy of the limbs in Fig. 4c is stronger than three other pictures. It is more like a triangle wave while others are more like sine wave. The echo of the limbs in Figs. 4a, b and d are very similar with each other except small differences in magnitude and energy. It is very difficult for us to distinguish them with traditional supervised learning techniques relying on hand-designed features.

### 3 Deep convolutional neural networks

#### 3.1 Theory of the DCNN

Deep learning is a branch of machine learning, which can be an alternative concept of artificial intelligence in many cases. It usually adopts DCNN structure, which is one of the most successful deep learning algorithms. It is a kind of multilayer supervising learning neural network [22]. The key components of DCNNs are the convolution and pooling in the hidden layer. In the feature extraction part of the network, convolution and pooling will be implemented alternatively. Multiple convolution filters work in parallel on input data to get the feature maps in convolution layer followed by pooling layer. Fig. 5a shows the schematic of the convolution filter and the pooling operation. The layers after

feature extraction part of this network are full connection layers, which includes logistic regression classifier. The input of the full connection is the output of the last pooling layer while the output of the full connection is classification results.

In our work, we choose softmax regression to classify the characteristics. Softmax is developed from logistic regression in order to solve multi-class problems. The function of softmax can be expressed as

$$h_{\theta}(x) = \begin{cases} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{cases} \quad (3)$$

$p(y^{(i)} = j|x^{(i)}; \theta)$  represents the probability of the input  $x^{(i)}$  of the  $i$ th sample belonging to the category  $j$ . The loss function of the softmax classifier can be expressed as

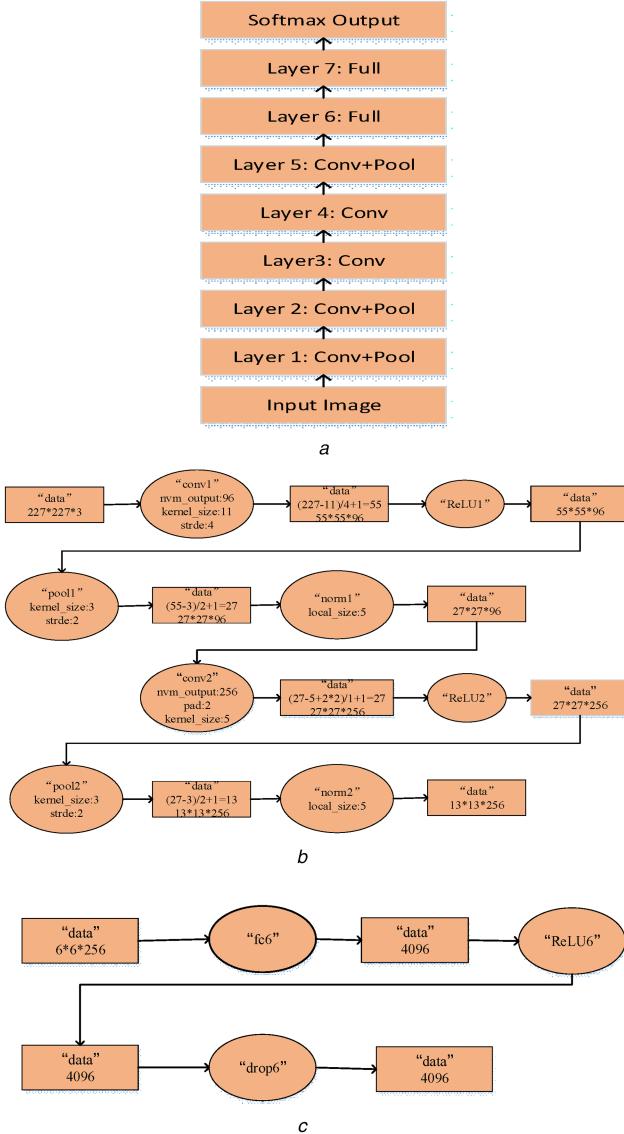
$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y_i = j\} \log \frac{e^{\theta_j^T x_i}}{\sum_{l=1}^k e^{\theta_l^T x_i}} \right] \quad (4)$$

Then the stochastic gradient descent is used to minimise the regulation of loss function  $J(\theta)$  in back propagation until the network converges or reaches the maximum iteration number. To prevent overfitting effectively, dropout is widely used. The nerve cell that is in the state of dropout would not participate in the forward propagation nor in the back propagation. In this way, the neural network is like trying a new structure for every input sample, which reduces the complex interrelationship of neurons. A simple DCNNs architecture with two convolution layers and one fully connected layer is shown in Fig. 5b.

#### 3.2 Specific network architecture

We employed Caffe [25], i.e. convolutional architecture for fast feature embedding, which is open-source and speeded up by the NVIDIA GPU and CUDA library, as the platform to analyse the spectrograms. The GPU that we used was the NVIDIA Quadro M4000. The spectrograms were used as the input of the DCNNs. Then, we transformed human identification problem into image recognition problem. Eighty per cent of the augmented spectrograms of each target were used as training data and the rest of the data were used to test.

We adopted AlexNet in which there were five convolution layers, two fully connected layers with 4096 hidden nodes in the first fully connected layer and an output layer as shown in Fig. 6a. The size of the input spectrograms of the whole AlexNet must be  $256 \times 256$ . We used the tools of Caffe to normalise the spectrograms directly. Fig. 6 shows partial internal structure of the network. Fig. 6b shows the first two convolutional layers and pooling layers while Fig. 6c shows the first full connection layer. Rectified linear units was used as activation function and followed by max pooling in each layer. We fine-tuned the network according to our experiments. In the configuration file of Caffe, the data of learning rate was adjusted to 0.0001 because when the learning rate was larger the loss would not converge and when the learning rate was smaller the process would be time consuming. In our experiments, the data of maximum iteration number was changed to 10,000 for stable results. The learning rate was reduced to  $0.0001 * 0.9 \wedge (\text{floor}(10000/5000))$  after every 5000 iteration, since we employed gradient descent method to solve the optimisation problem. The weight decay was changed to 0.0005. The definition of the network for training and validation as well as the parameters for every layer are recorded in a particular file. In this file, the spectrograms are resized to  $227 \times 227$ , since the size of the input spectrograms of the first convolution layer is  $227 \times 227$ . The batch sizes in ‘train’ part and ‘test’ part were adjusted to 32 and 16, respectively. Batch size represents the number of samples taken in each iteration. The average of the gradient of these samples is used to update the parameters of the network. The batch size determines the direction of gradient



**Fig. 6** Framework of Alexnet

(a) Structure diagram of AlexNet, (b) First two convolution layers and pooling layers, (c) First full connection layer

descent and the effect and rate of convergence, as well as memory utilisation. As the spectrograms used to train were more than that were used to test, the batch size in ‘train’ part was more than that in ‘test’ part. In the full connected layer, we adapted the learning rate of bias to 10 and the learning rate of weight to 20 to speed up the learning rate in this layer.

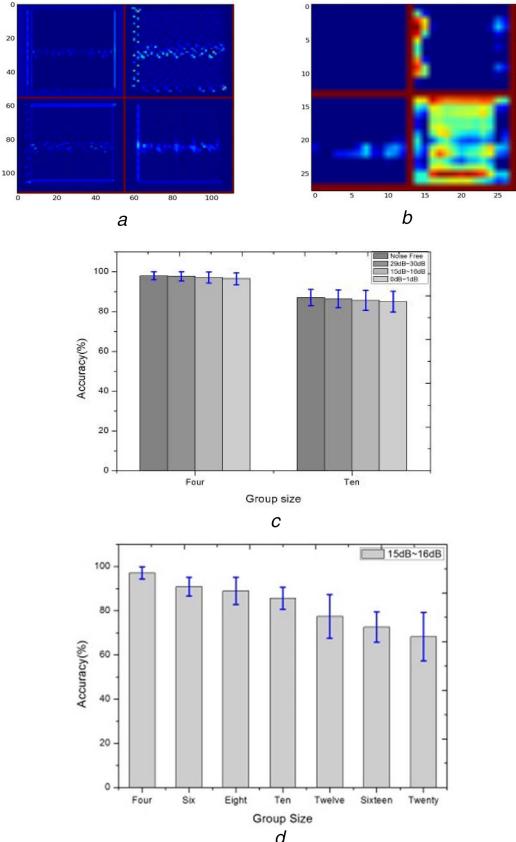
After 10,000 iterations, which took about 15 min for four subjects while 30 min for ten subjects, our own network was generated which stored some model parameters like weight. Then we chose several spectrograms randomly to test our network. The testing progress took about 5 s which could be very suitable for real-time human identification application.

Figs. 7a and b show the feature maps of the first convolution layer and the third convolution layer, respectively. Compare to Fig. 7a, in Fig. 7b there is already a little abstract for us to obtain the physical insights into our case. In the future, we plan to study the learned features of the DCNNs, which will help us get better insights.

## 4 Results and discussion

### 4.1 Effect of noise

In this sub-part, our own network was employed to detect the noise immunity of the algorithm. We added different grades (SNR = 29 – 30 dB, 15 – 6 dB, 0 – 1 dB) of random noise to the echo, respectively. For example, we added random noise



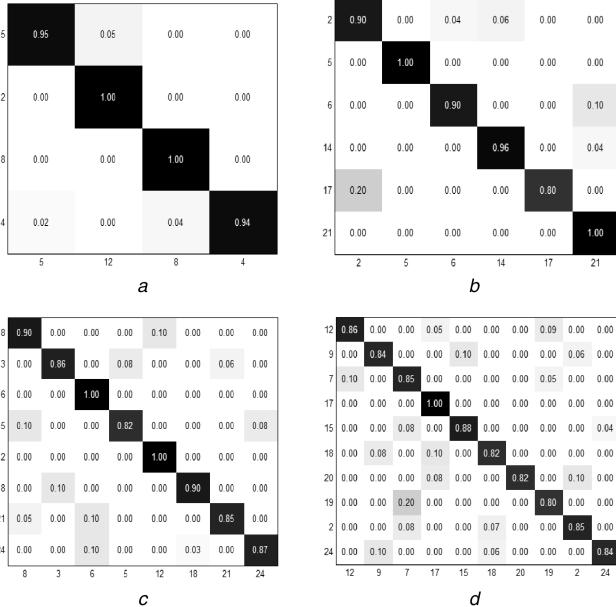
**Fig. 7** Feature maps and impacts of noise and group sizes

(a) Feature maps of first convolution layer, (b) Feature maps of the third convolution layer, (c) Impact of noise, (d) Impact of group size

(SNR = 0 – 1 dB) to the raw signals that we gathered in Section 2. We discussed the results of two different group sizes (4 and 10). Each group was a combination of different signals from the corresponding number of targets. The subjects of each group were chosen randomly from all subjects, and after that, we randomly chose 100 spectrograms of each target at each noise grade to test the noise immunity of the algorithm. The whole process was repeated for 50 times. The results of the groups with the same number of subjects at the same noise grade varied slightly and the average results were taken as the terminal results. Fig. 7c shows the classification results of the two groups. We can observe that the classification performance can be compromised when the noise increases in both group sizes. However, compared to the extent of noise growth the decrease in accuracy is very small, which can be ignored, that demonstrates the algorithm has good anti-noise performance.

### 4.2 Effect of group size

In this sub-section, we still used our own network. We assessed the performance of our system under the scenario where we identified an individual from a group of  $N$  people with the random noise (SNR = 15 – 16 dB) in the spectrograms. The groups were combinations of different signals from 4 people, 6 people, 8 people, 10 people, 12 people, 16 people and 20 people, respectively. The subjects of each group were chosen randomly from all subjects and in each case, we randomly selected 100 spectrograms of each subject to test the effect of group size. The whole process was repeated for 50 times. The results of the groups with the same number of subjects varied slightly and the average results were taken as the terminal results which are shown in Fig. 7d. We can observe that the accuracy declines with the group size increasing. Even so, the accuracy can reach ~85% or higher when the number of the group is <10, which can compare favourably with the method based on vision (like camera) or biological feature (like fingerprints).



**Fig. 8** Confusion matrices of four group sizes (4, 6, 8, 10). The x- and y-axis indicate the serial number of the subjects

Figs. 8 and 9 show the confusion matrices of random cases for the seven different group sizes. Since the results of the groups with the same number of subjects varied slightly, result of one case is similar to the average result. Fig. 9a shows the accuracy as well as the loss of training and testing process of a random case for six people. The loss showed is the value of the loss function in formula (4). As it converges to a number tending to zero, the network does not overfit. Fig. 8 shows that individuals can be identified from a group with high possibility when the number of the group is <10.

#### 4.3 Comparison with support vector machine (SVM) and naive Bayes (NB)

We also employed traditional supervised learning techniques (SVM and NB) to identify individuals with a group of four people. The features we chose to extract were listed as follows:

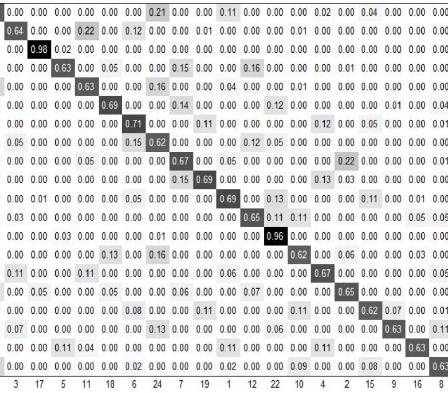
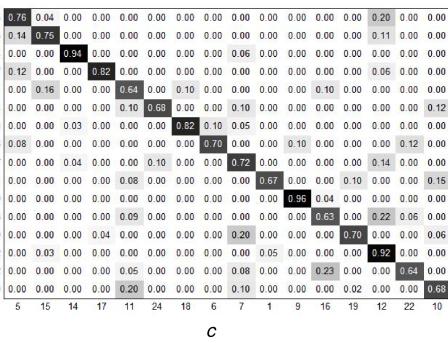
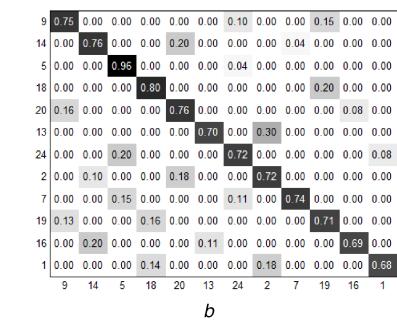
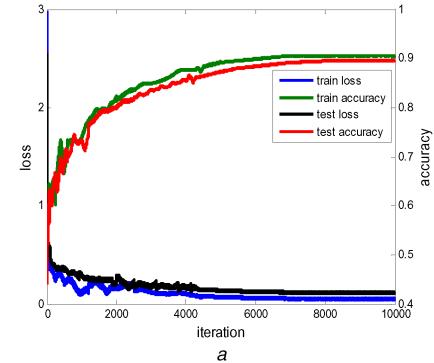
- (i) Torso/body radial velocity;
- (ii) Offset of total Doppler;
- (iii) Total bandwidth of Doppler;
- (iv) Total energy of Doppler;
- (v) Micro-Doppler cycle;
- (vi) Principal component analysis (PCA) based features.

We chose two features extracted from PCA on spectrograms. The first feature was referred to as the latent feature, and was the principal component variance of the covariance matrix. While the second feature that we utilised was Hotelling's T-squared statistics.

Fig. 10a shows the noise immunity of DCNN, SVM and NB, respectively. It is apparent that the noise immunity of SVM and NB cannot compare with that of the DCNN. Figs. 10b and c show the confusion matrices of SVM and NB, respectively, and the average accuracy is 84.25 and 78%, which still cannot compare with the accuracy of the DCNNs. Meanwhile the traditional supervised learning techniques rely severely on the extracted features, which means different features will lead to different results, and require domain knowledge of each problem. Furthermore, the process of choosing and extracting features is very complex. These factors limit their use by people who are not familiar with related fields. However neither of them is a problem for DCNNs, which means DCNNs is easier to be used widely in real life.

#### 4.4 Comparison with other techniques

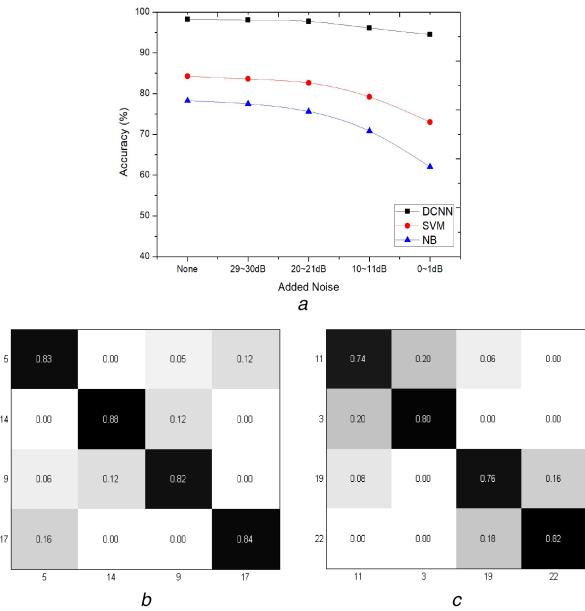
In the introduction, the benefits and drawbacks of various methods, which have been explored to recognise persons, have been demonstrated. In this sub-section, we mainly present the difference



**Fig. 9** Results of different group size

(a) Accuracy and loss of a random case for six people, (b), (c), (d) Confusion matrices of three group sizes (12, 16, 20). The x- and y-axis indicate the serial number of the subjects

of accuracy between other methods and our method. Since other methods, like image recognition, fingerprints identification, do not belong to the radar category and we are not familiar with them, to realise these methods is very tough for us. Considering that the published papers usually present relatively great results, we choose to compare our results with theirs directly. Ng *et al.* [7] proposed an approach to identify human based on extracted gait features using images. The maximum group size of this approach is 11 and its accuracy is 84%, which is close to our results. In [11], WIFI signal was employed to sense human body for the first time. The maximum group size of this approach is 6 and its average accuracy is 93–77% from a group of two to six people, which shows that our approach is better considering the number of subjects as well as the



**Fig. 10** Results of NB and comparison between different classifiers

(a) Impact of noise for DCNN, SVM and NB, (b) Confusion matrices SVM, (c) Confusion matrices NB

accuracy of classification. Footstep induced structural vibration is employed in indoor person identification with the maximum group size of 5 [26]. The accuracy for five people is 70% which is far less than our method. The groups mentioned above are combinations of different signals from the corresponding number of targets. After these comparisons, we can find our results are better or no lower than others.

## 5 Conclusions

In this paper, we propose Radar-ID a radar-based human identification system. In this system, we apply the DCNNs to radar micro-Doppler signatures for personal recognition without any feature extraction pre-processing which will have great advantages in practical applications. By using the DCNNs, human beings can be identified successfully. When we identify an individual from a group of  $N$  people, the accuracy can reach more than 85% when the group size is  $<10$ . To be specific, the average accuracy is about 97.1% for 4 people, 90.9% for 6 people, 89.1% for 8 people, 85.6% for 10 people, 77.4% for 12 people, 72.6% for 16 people and 68.9% for 20 people, respectively. Also, this algorithm has excellent anti-noise performance. By contrast, the performance of SVM and NB, the typical traditional supervised learning techniques, are far from satisfactory in both aspects. We believe that it shows the potential of the DCNNs for human identification problems based on radar micro-Doppler signatures.

While our results are satisfying, this method still has some limitations to be solved for extensive applications. First, since the shape of the micro-Doppler signature is the key for classification, when the spectrograms of two people are similar or the motions are irregular the accuracy may decline. Second, the condition we considered in the experiments, where the participants ran along the line-of-sight path of the radar, is simple compared with actual realistic running pattern. In the future, we will do further researches about non-LOS scenarios to promote the use of our system. In addition, in our experiments we identify one person from a maximum group size of 20 people. To apply our system to the practical conditions we also need to identify a person who does not belong to the sample group in the future and expand the maximum group size.

## 6 Acknowledgments

This work was supported by the National Natural Science Foundation of China under grant 6150010825, the Fundamental

Research Funds for the Central Universities under grant NS2016040, the Fundamental Research Funds for the Central Universities under grant NJ20150020 and the China Scholarship Council 201606835062.

## 7 References

- [1] Kumar, A., Zhou, Y.B.: 'Human identification using finger images', *IEEE Trans. Image Process.*, 2012, **21**, (4), pp. 2228–2244
- [2] Li, H.Y., Sun, Y.J.: 'Computer-generated image identification based on image gradient'. Int. Conf. on Intelligent Systems Research and Mechatronics Engineering (ISRME), Zhengzhou, China, April 2015, pp. 394–397
- [3] DeLoney, C.: 'Person identification and gender recognition from footstep sound using modulation analysis'. ISR Technical report, 2008
- [4] Wang, C., Zhang, J.P., Wang, L., et al.: 'Human identification using temporal information preserving gait template', *IEEE Trans. Pattern Anal. Mach. Intell.*, 2012, **34**, (11), pp. 2164–2176
- [5] Krucki, K., Asari, V., Borel, C., et al.: 'Human re-identification in multi-camera systems'. IEEE Applied Imagery Pattern Recognition Workshop, Washington, DC, USA, October 2014, pp. 1–7
- [6] Geiger, J.T., Kneißl, M., Schuller, B.W., et al.: 'Acoustic gait-based person identification using hidden markov models'. Proc. of the 2014 Workshop on Mapping Personality Traits Challenge and Workshop, Istanbul, Turkey, 2014, pp. 25–30
- [7] Ng, H., Ton, H.-L., Tan, W.-H., et al.: 'Human identification based on extracted gait features', *Int. J. New Comput. Archit. their Appl. (IJNCAI)*, 2011, **1**, pp. 358–370
- [8] Zhou, Z., Du, E.Y., Thomas, N.L., et al.: 'A new human identification method: sclera recognition', *IEEE Trans. Syst. Man and Cybernetics Part a-Syst. Humans*, 2012, **42**, (3), pp. 571–583
- [9] Ikeda, T., Ishiguro, H., Miyashita, T., et al.: 'Pedestrian identification by associating wearable and environmental sensors based on phase dependent correlation of human walking', *J. Ambient Intell. Humanized Comput.*, 2014, **5**, (5), pp. 645–654
- [10] Falguera, F.P.S., Marana, A.N., Falguera, J.R.: 'Fusion of fingerprint recognition methods for robust human identification'. IEEE Int. Conf. on Computational Science and Engineering, Sao Paulo, Brazil, July 2008, pp. 413–420
- [11] Zhang, J., Bo, W.: 'WiFi-ID: human identification using WiFi signal'. IEEE Int. Conf. on Distributed Computing in Sensor Systems, Washington, DC, USA, May 2016, pp. 75–82
- [12] Yang, Y., Lu, C.: 'Human identifications using micro-Doppler signatures'. Int. Conf. on Antennas, Radar and Wave Propagation, Baltimore, Maryland, April 2008, pp. 69–73
- [13] Tahmoush, D.: 'Review of micro-Doppler signatures', *IET Radar, Sonar Navigat.*, 2015, **9**, (9), pp. 1140–1146
- [14] Seyfoglu, M.S., Özbayoglu, A.M., Gurbuz, S.Z.: 'Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities', *IEEE Trans. Aerosp. Electron. Syst.*, 2018, **PP**, (99), p. 1
- [15] Ricci, R., Balleri, A.: 'Recognition of humans based on radar micro-Doppler shape spectrum features', *IET Radar Sonar Navig.*, 2015, **9**, (9), pp. 1216–1223
- [16] Fioranelli, F., Ritchie, M., Griffiths, H.: 'Performance analysis of centroid and SVD features for personnel recognition using multistatic micro-Doppler', *IEEE Geosci. Remote Sens. Lett.*, 2016, **13**, (5), pp. 725–729
- [17] Gürbüz, S., Erol, B., Cagliyan, B., et al.: 'Operational assessment and adaptive selection of micro-Doppler features', *IET Radar, Sonar Navigat.*, 2015, **9**, (9), pp. 1196–1204
- [18] Banka, A.A., Ajaz, D.: 'Human identification based on gait'. Int. Conf. on Biometrics Technology, At PSG College of Technology, Coimbatore, India, May 2010
- [19] Hinton, G., Deng, L., Dong, Y., et al.: 'Deep neural networks for acoustic modeling in speech recognition', *IEEE Signal Process. Mag.*, 2012, **29**, (6), pp. 82–97
- [20] Krizhevsky, A., Sutskever, I., Hinton, G.: 'Imagenet classification with deep convolutional neural networks'. Int. Conf. on Neural Information Processing Systems, Lake Tahoe, NV, USA, 2012, vol. 25, no. 2, pp. 1097–1105
- [21] Mikolov, T., Deoras, A., Povey, D., et al.: 'Strategies for training large scale neural network language models'. IEEE Workshop on Automatic Speech Recognition & Understanding (ASRU), Waikoloa, HI, USA, December 2011, pp. 196–201
- [22] Kim, Y., Moon, T.: 'Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks', *IEEE Geosci. Remote Sens. Lett.*, 2016, **13**, (1), pp. 8–12
- [23] Javier, R., Kim, Y.: 'Application of linear predictive coding for human activity classification based on micro-Doppler signatures', *IEEE Geosci. Remote Sens. Lett.*, 2013, **11**, (10), pp. 781–785
- [24] Chen, V., Ling, H.: 'Time-frequency transforms for radar imaging and signal analysis' (Artech House, Norwood, MA, USA, 2002)
- [25] Jia, Y., Shelhamer, E., Donahue, J., et al.: 'Caffe: convolutional architecture for fast feature embed-ding'. the 22nd ACM int. Conf. on Multimedia, Orlando, Florida, USA, November 2014, pp. 675–678
- [26] Pan, S., Wang, N., Qian, Y., et al.: 'Indoor person identification through footprint induced'. Int. Workshop on Mobile Computing Systems & Applications, Santa Fe, NM, USA, 2015, pp. 81–86