Personnel Recognition and Gait Classification Based on Multistatic Micro-Doppler Signatures Using Deep Convolutional Neural Networks

Zhaoxi Chen, Gang Li, *Senior Member*, *IEEE*, Francesco Fioranelli, *Member*, *IEEE*, and Hugh Griffiths, *Fellow*, *IEEE*

***Abstract*—In this letter we propose two methods for personnel recognition and gait classification using deep convolutional neural networks (DCNNs) based on multistatic radar micro-Doppler signatures. Previous DCNN based schemes have mainly focused on monostatic scenarios, whereas directional diversity offered by multistatic radar is exploited in our work to improve classification accuracy. We first propose the voted monostatic DCNN method (VMo-DCNN), which trains DCNNs on each receiver node separately, and fuses the results by binary voting. By merging the fusion step into the network architecture, we further propose the multistatic DCNN method (Mul-DCNN), which performs slightly better than VMo-DCNN. These methods are validated on real data measured with a 2.4 GHz multistatic radar system. Experimental results show that Mul-DCNN achieves over 99% accuracy in armed/unarmed gait classification using only 20% training data and similar performance in two-class personnel recognition using 50% training data, which are higher than the accuracy obtained by performing DCNN on a single radar node.**

***Index Terms*—Convolutional neural networks, data fusion, deep learning, micro-Doppler, multistatic radar, target classification.**

I. Introduction

MICRO-Doppler refers to the additional Doppler frequency shift of moving targets generated by vibration, rotation, etc. with respect to their main Doppler component [1]. Target classification using micro-Doppler signatures has seen a rapid growth in recent years [2-6], with application in fields including surveillance [2][3], healthcare [4][5] and human-computer interaction [6]. Based on the human micro-Doppler signature, personnel recognition and human activity classification have attracted much attention [5, 7-10]. In [7], empirical features with clear physical meaning are used to train a support vector machine (SVM) classifier. Similar classification tools are used in [8] on dual frequency radar micro-Doppler signatures. The authors of [9] propose some features based on singular value decomposition (SVD) of the spectrogram, which yield good performance in classification of unarmed/armed personnel outdoors. Principal component analysis (PCA) is used in [5] for feature extraction, and a more robust tool L1-PCA is utilized in [10] for indoor human limb motion classification.

The newly developed deep learning algorithms have been introduced into radar target classification. One of the preliminary works by Kim [11] investigated the feasibility of using deep convolutional neural network (DCNN) in micro-Doppler based classification tasks. The authors used a DCNN with a straightforward structure to distinguish human from three classes of non-human objects and to classify six classes of human activities. In [12], a similar method was used for hand gesture classification. More sophisticated DCNN architectures were used later, including 7-layer DCNN [13], transfer learned AlexNet and VGG-16 network [14] and a three-layer semi-supervised auto-encoder [15]. New problems such as low latency classification [16] and multi-target human gait classification [9, 17] have also been taken into consideration.

It is well known that micro-Doppler signatures depend on the aspect angle between the target movement and the radar line of sight. Classification performance suffers severe degradation when the aspect angle is close to 90°, but it degrades slightly at smaller aspect angles, e.g. 30° [7]. Because multistatic radar observes targets from different lines of sight, it has the potential to alleviate the negative effect of large aspect angles and hence to improve the classification accuracy by using proper multi-view fusion methods. One of the pioneer works [18] uses a fused spectrogram from multistatic radar data, but the algorithm is tested only on synthetic data generated by video motion capture. In further study [9] and [19], real data are collected by the multistatic radar system NetRAD for the classification of armed/unarmed personnel targets. Using empirical features and off-the-shelf classifiers, the authors train unique classifiers for each receiver node, and then fuse the classification results by binary voting. The fused result shows improved accuracy compared to each receiver node itself. Another approach to classification with multistatic radar firstly fuses features from different nodes and then feed them into classifiers. Ref. [20] uses brute force search and other less computationally intensive algorithms, e.g. T-test and mutual information criteria, to find three optimal/sub-optimal features out of twelve pre-defined features at each receiver node. Instead of feature selection, Ref. [21] uses linear combination, i.e. PCA, to fuse features obtained from 4×4 MIMO channels.

A natural thought is to combine the advantage of DCNN and multi-view fusion. A novel work [16] investigates the feasibility of combining data from different aspect angles to

This work was supported by the National Natural Science Foundation of China under Grants 61661130158, by Shenzhen Fundamental Research Program, by the Royal Society Newton Advanced Fellowship,, and by the IET A. F. Harvey Prize awarded to Hugh Griffiths in 2013 and the Engineering and Physical Sciences Research Council [EP/G037264/1]. Corresponding author: Gang Li. Email: gangli@tsinghua.edu.cn

Z. Chen and G. Li are with the Department of Electronic Engineering, Tsinghua University, Beijing 100084, China. G. Li is also with the Research Instituete of Tsinghua University in Shenzhen, Shenzhen, 518057, China.

F. Fioranelli is with the University of Glasgow, Glasgow G12 8QQ, UK.

H. Griffiths is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, UK.

improve classification accuracy, using deep learning and boosting trees. However, the data are measured by monostatic radar at different aspect angles rather than by a multistatic radar system simultaneously. To the best of our knowledge, classification based on multistatic micro-Doppler signatures using DCNN has not been sufficiently investigated.

In this letter, performing DCNN on a single radar node is called monostatic DCNN method (Mo-DCNN). With the data collected by three radar nodes of a multistatic radar, we propose two DCNN based methods for personnel recognition and gait classification. One is the voted monostatic DCNN method (VMo-DCNN), in which we fuse the results of Mo-DCNN of each node via binary voting. The other one is the multistatic DCNN method (Mul-DCNN), in which a fusion layer is added to the network thus the fusion step is conducted inside the DCNN automatically. Due to the existence of the fusion layer, Mul-DCNN is able to learn fusion rules automatically and has the potential to achieve better performance. Both VMo-DCNN and Mul-DCNN are tested on real data and show significant accuracy improvement over Mo-DCNN for both personnel recognition and gait classification tasks.

The remaining parts of this letter are organized as follows. Section II describes the multistatic radar dataset. Section III demonstrates the DCNN architecture and the process of DCNN training. Section IV presents the results of the two classification tasks. Finally, we conclude the letter in Section V.

II. Multi-static Radar Dataset

The data used in this letter were collected in July 2015 by NetRAD, a coherent multistatic pulsed radar system developed at University College London [9]. NetRAD consists of three nodes deployed on a linear baseline and operating at 2.4 GHz, with linear up-chirp modulation, 45 MHz bandwidth and 5 kHz pulse repetition frequency (PRF). Fig. 1 shows the geometry of the experimental scene. All antennas of the three nodes are pointing at zone 5. The node in the middle (node 1) is a transmitter-receiver node, whereas the other nodes (node 2 and 3) on both sides are receivers only. Thus, three-channel synchronized data can be collected simultaneously. Each set of data was recorded for 5 s, during which a single person walked towards the baseline in one of the six zones, either moving his arms freely (referred to as "unarmed" case) or holding a metallic pole (referred to as "armed" case). The whole dataset consists of 2 persons, 2 actions (i.e. armed and unarmed), 6 zones and 5 repetitions for each case, making a total number of 120 three-channel recordings. We further duplicate the size of the dataset by splitting every piece of data into two pieces both with 2.5 s duration, and discard the data collected in zone 5 due to some missing data. In summary, we use a dataset containing 200 samples collected by the three nodes.

Typical data samples collected by node 1 are visualized in Fig.2 using the Short Time Fourier Transform (STFT). The main Doppler component is about 20 Hz, corresponding to the torso speed of the subject. The other components mainly indicate the movement of limbs. It is easy to find differences between the armed (Fig.2 (b) and (d)) and unarmed (Fig.2 (a) and (c)) gaits. In the unarmed case, higher Doppler bandwidth is observed due to freely swinging arms. In the armed case, the movement of arms is restricted by the metallic pole in hands,

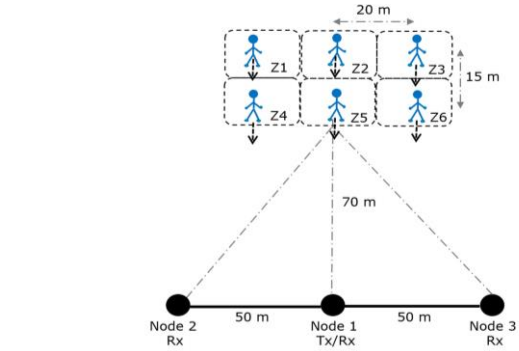

Fig. 1. The NetRAD radar system setup [9]

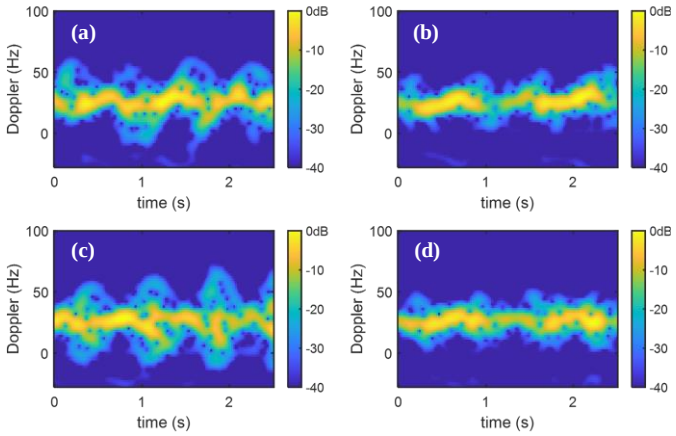

Fig. 2. Typical spectrograms of different person and activity (node 1): (a) person A, unarmed; (b) person A, armed; (c) person B, unarmed; (d) person B, armed

resulting in a more condensed micro-Doppler pattern. The difference between the persons A (Fig.2 (a) and (b)) and B (Fig.2 (c) and (d)), however, is less noticeable. When both persons are unarmed (Fig.2 (a) and (c)), the gait of person A shows slight asymmetry. The difference is even harder to find by human eyes when they are both armed (Fig.2 (b) and (d)).

In this letter, we focus on the tasks of gait classification and personnel recognition. Both are two-class classification tasks. In the gait classification task, the two classes are armed and unarmed gaits, regardless of the person. In the personnel recognition task, we try to classify persons A and B regardless of whether they are armed or unarmed. The personnel recognition task is much more challenging.

III. DCNN Implementation and Training

A. Data Pre-processing for Transfer Learning

Here the transfer learned DCNN is used to distinguish different gaits and different personnel targets. Transfer learning [22] is generally a technique that aims to transfer the knowledge learned from one task to another related but different one. In the field of DCNN, it refers to utilizing the information of a network pre-trained on a large dataset to train a different network on a small dataset, which has been successfully used in the design of many DCNNs [23, 24] to alleviate overfitting problems. It can be done in two steps: (1) replace the last few layers in the pre-trained network by new designed ones and initialize them randomly; (2) train on the small dataset (referred

to as "fine-tuning"). Recent work has explored the feasibility of using CNN pre-trained on optical image dataset to fine-tune micro-Doppler spectrograms [14]. However, the time-frequency spectrograms have only one channel, whereas optical images typically have RGB channels. The authors of [14] simply copy STFT spectrograms for the three input channels to solve this dimension mismatch problem, which is equivalent to regarding the spectrograms as grayscale images.

In our method, the STFT spectrograms with three different window sizes are used as different channels of input data. Inspired by multi-resolution analysis methods such as wavelet, we believe that the magnitude of spectrograms in different time-frequency resolutions provide richer information than single resolution ones. Specifically, we first calculate log-scale spectrograms with dimension 128 (frequency) × 125 (time) using a Blackman window and a threshold of -40dB. Then the spectrograms are normalized to the interval [0, 255] to match the range of optical images, and finally the mean value of each spectrogram is subtracted. We set the window size to 0.13 s, 0.26 s, and 0.51 s, respectively. One may refer to Fig. 3 to find an example of the input time-frequency spectrograms. Fig.3 (b) to (d) demostrate the spectrograms with increasing window size, corresbonding to increasing frequency resolution and decreasing time resolution. In Fig. 3(a), spectrograms of different window sizes are stacked as RGB image compoments, which are used exactly as the inputs of DCNN. The R, G and B channels of Fig.3 (a) are identical to those of Fig.3(b), (c) and (d), respectively.

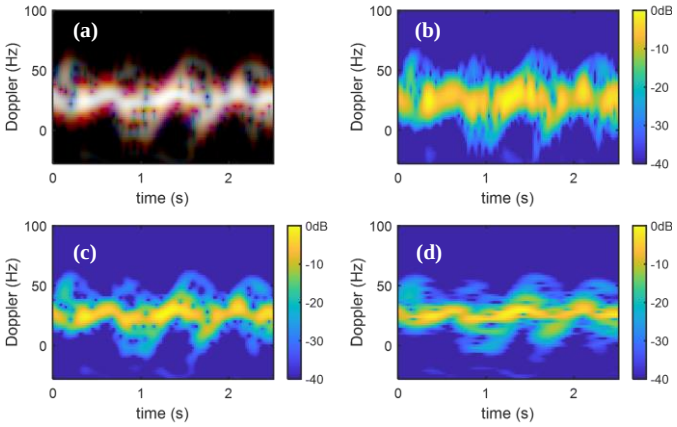

Fig. 3. Three-channel spectrogram as input (node 1): (a) Spectrograms stacked as an RGB image; (b) – (d) Spectrograms in default MATLAB colors, FFT window size equals to 0.13 s, 0.26 s, 0.51 s, respectively

B. *DCNN Architecture*

Using data from a single node as input, a six-layer DCNN is carried out for Mo-DCNN (shown in Fig.4, where the number follows '#' indicates the feature depth). The Rectified Linear Unit (ReLU) activation is used after each layer except the last layer fc 6, where softmax is used. The idea for this architecture is straightforward. Optical images and spectrograms share some low level features, e.g. edges and curves, which are captured in the first several convolution layers of a network. Therefore, the first three layers (conv 1-3) are identical to and initialized with the first three convolution layers of pre-trained VGG-f network [25], which is an eight-layer DCNN architecture originally used for optical image classification. Here VGG-f is used just for example, and one may change to other network architectures with corresponding modification. We add the subsequent convolution layer (conv 4) to reduce dimensionality of the feature map along Doppler axis, resulting in an output with dimension $1 \times 7 \times 64$ (Doppler × time × depth). Now, the output of conv 4 could be considered as seven different feature vectors with length 64, each containing information of the spectrogram within different (but overlapped) time interval. Finally, these vectors are fed into another two fully connected layers (fc 5-6) followed by a softmax activation to produce the final output. The output, a 7×2 matrix, represents the estimated Bernoulli distribution of the two classes within seven different time interval. When conducting back propagation, we treat the above-mentioned feature vectors as separate training samples. However, when testing, we mean-pool the output (dimension 7×2) along the timeline and get an averaged probability distribution (dimension 1×2), which improves classification robustness significantly.

The proposed VMo-DCNN simply fuses the Mo-DCNN output of three nodes by binary voting. For the Mul-DCNN (shown in Fig.5, where some details are omitted for a clearer view), the first four layers in Mo-DCNN are copied for each receiver nodes. The corresponding layers share the same weights except for layer 4. Weight sharing reduces the total number of parameters, thus mitigating potential overfitting problem. Here we add a fusion layer to aggregate information from three branches corresponding to the three radar nodes. In this layer, the three output feature maps of layer 4 are max-pooled along the node dimension (elementwise maximum operation). The subsequent layers are identical to those in Mo-DCNN.

For training speed-up and overfitting prevention, batch

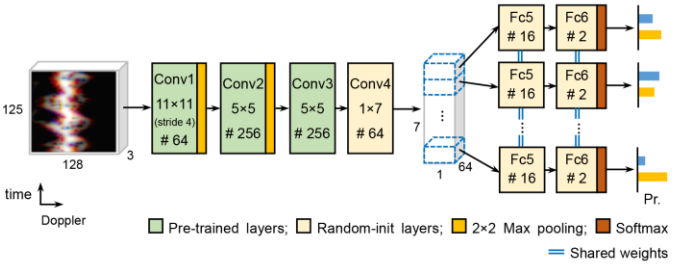

Fig. 4. The DCNN architecture for Mo-DCNN (1 node per network as input)

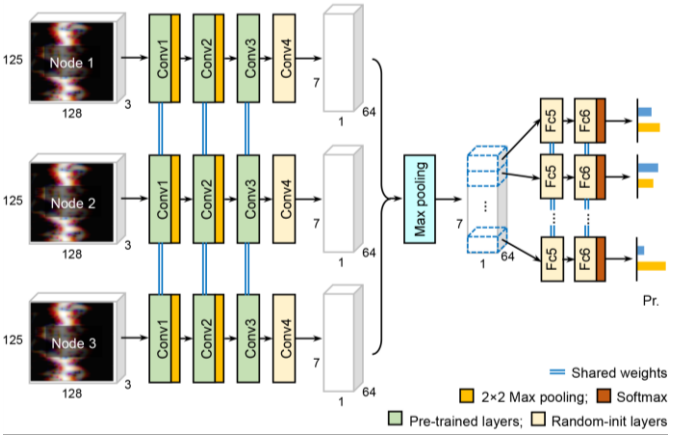

Fig. 5. Proposed DCNN architecture for Mul-DCNN (3 nodes as input together)

normalization [26] and dropout [27] are used in the DCNN architecture except for the last two layers, since these two layers contain very small number of weights and are not likely to become overfitted.

C. Training details

We implement the proposed neural network using MatConvNet [28], a MATLAB based open source CNN toolkit. For training, we use the Adam solver [29] with parameters $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1 \times 10^{-8}$ and a fixed batch size of 5, with mild complex Gaussian noise added to raw training samples to alleviate overfitting problems. Training lasts for 300 epochs in total. For the first 100 epochs, we initialize the first three convolution layers with VGG-f net [25] pre-trained on ImageNet dataset and set their learning rates to zero. The rest layers are initialized randomly and trained with learning rate $\alpha = 5 \times 10^{-3}$ (for 20% and 33% training data ratios) or $\alpha = 2 \times 10^{-3}$ (for 50% training data ratio). Then we set small (0.1α) learning rate for the first three layers while keeping the others unchanged for another 100 epochs. Finally, all learning rates are reduced to a tenth for additional 100 epochs, making the network converge. The training is carried out on an NVIDIA GTX1060 GPU with 6 GB memory. The training process takes minutes for 300 epochs, while testing takes only a few milliseconds per sample, which is affordable in some real-time scenarios. We summarize running times of training and testing in Fig. 6. Note that training time increases with the number of training samples while testing time only relies on the network architecture.

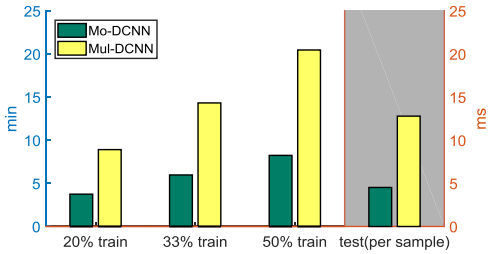

Fig. 6. Running times of training (left axis) and testing (right axis)

IV. Experimental Results

A. Randomly Partitioned Training Set

We first investigate the performance under three different ratios of training data, i.e. 20%, 33%, 50%. In the case of 20% training ratio, we partition the dataset into five folds randomly and evenly and conduct 5-fold cross validation using one fold as training set at each time. Since the two samples from the same piece of data tend to be similar, we put them either into training set or testing set. To make the result even more robust, the 5-fold cross validation is repeated for three times with statistically independent dataset partitions, making a total 15 repetitions. Similarly, we conduct 3-fold cross validation for five times in the case of 33% training ratio (15 repetitions in total) and 2-fold cross validation for eight times in the case of 50% training ratio (16 repetitions in total), resulting in almost the same total repetitions for all three ratios.

The performances of both tasks, i.e., gait classification and personnel target identification, are validated using the above-mentioned partition setups. Tables I and II show the minimum, maximum, and average accuracy among 15 or 16 independent trials. Columns 1 to 3 represent the accuracy of Mo-DCNN, namely the results of monostatic DCNN at each single node. Columns 4 to 5 represent the results of VMo-DCNN and Mul-DCNN, respectively. As shown in Table I, the accuracy of gait classification is high enough with 20% training data, in which Mul-DCNN performs the best in average classification accuracy. We also find that VMo-DCNN fails to outperform Mo-DCNN with node 1 data, but this minor performance gap could be neglected considering statistical variance. In the more challenging personnel recognition task (shown in Table II), we try all training ratios. It is observed that both VMo-DCNN and Mul-DCNN show significant accuracy improvement over Mo-DCNN. Mul-DCNN has the best overall performance again, though VMo-DCNN shows better minimum or maximum accuracy in certain scenarios. We pay additional attention to the minimum accuracy in both tasks, since the worst case indicates the robustness of an algorithm. We are glad to see that, in most scenarios, Mul-DCNN in both tasks improves the worst accuracy effectively compared to Mo-DCNN.

B. Training on One Zone

In addition to the random partitioning, we evaluate the proposed methods using one-zone data as the training set and the rest for testing. This experiment is more practical since the

TABLE I
RANDOM PARTITION: GAIT CLASSIFICATION ACCURACY (%)

		Rx node 1	Rx node 2	Rx node 3	VMo-DCNN	Mul-DCNN
20% training	min	**98.75**	95.62	94.37	98.12	**98.75**
	max	**100.00**	98.75	**100.00**	**100.00**	**100.00**
	average	99.50	97.17	98.00	99.33	**99.63**

TABLE II
RANDOM PARTITION: PERSONNEL RECOGNITION ACCURACY (%)

		Rx node 1	Rx node 2	Rx node 3	VMo-DCNN	Mul-DCNN
20% training	min	91.25	88.75	87.50	91.87	**93.12**
	max	98.75	97.50	96.87	**99.37**	**99.37**
	average	94.50	94.33	91.96	97.13	**97.42**
33% training	min	93.08	89.29	92.31	**96.15**	93.85
	max	99.23	98.46	99.23	**100.00**	**100.00**
	average	97.10	95.73	96.65	98.64	**98.98**
50% training	min	97.00	95.00	98.00	**99.00**	**99.00**
	max	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
	average	98.12	98.31	98.75	99.75	**99.94**

TABLE III
TRAIN ON ONE ZONE: ACCURACY (%)

Task	Training zone	Rx node 1	Rx node 2	Rx node 3	VMo-DCNN	Mul-DCNN
Gait classification	zone 1	98.75	93.00	98.75	99.75	**99.87**
	zone 2	99.12	93.37	98.25	**99.38**	99.37
	zone 3	**100.00**	89.87	99.88	**100.00**	**100.00**
	zone 4	96.37	88.75	97.62	97.38	**98.12**
	zone 6	99.13	93.87	97.75	**99.50**	99.00
	average	98.67	91.77	98.45	99.20	**99.27**
Personnel recognition	zone 1	94.13	81.75	83.00	90.88	**95.50**
	zone 2	87.75	80.50	82.12	**89.88**	89.62
	zone 3	**94.75**	85.62	89.25	93.62	93.25
	zone 4	78.37	85.63	86.87	89.00	**89.25**
	zone 6	95.00	95.12	89.88	96.13	**97.25**
	average	90.00	85.72	86.23	91.90	**92.97**

testing data have aspect angles that the classifier has never seen in training, which is often the case in the real applications. We perform experiments on both tasks, i.e. gait classification and personnel target identification. Five repetitions are done for each deterministic data partition and each task. The average performances of different methods are provided in Table III. Compared to random partitioning with 20% training data, the training ratio is the same but the accuracy falls as expected. However, in both tasks, VMo-DCNN and Mul-DCNN outperform all single nodes in nearly all training zones. Moreover, Mul-DCNN performs the best in terms of the average accuracy in most zones. This result indicates the robustness of the proposed methods.

V. Conclusion and discussion

In this letter, we combined the superiority of DCNN and multistatic radar in micro-Doppler signature classification. The novel architecture design in the proposed Mul-DCNN enables data fusion within the DCNN, which outperforms processing at a single node (Mo-DCNN) as well as binary voting of multiple nodes (VMo-DCNN). To fully utilize the three channels in pre-trained DCNN, we proposed a novel pre-processing technique using multi-window-size spectrograms as input to the network. Experiments on real data show that Mul-DCNN achieves over 99% accuracy in gait classification using only 20% training data and similar performance in personnel recognition using 50% training data. Future work will aim to collect and analyze more data of different subjects and different classes of activities to test the proposed DCNN architectures.

References

[1] V. C. Chen, "Doppler signatures of radar backscattering from objects with micro-motions," *IET Signal Process.*, vol. 2, no. 3, pp. 291–300, Sep. 2008.

[2] B. K. Kim, H. S. Kang and S. O. Park, "Drone classification using convolutional neural networks with merged Doppler images," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 38-42, Jan. 2017.

[3] R. Zhang, G. Li, C. Clemente, and J. J. Soraghan, "Multi-aspect micro-Doppler signatures for attitude-independent L/N quotient estimation and its application to helicopter classification," *IET Radar, Sonar Navigat.*, vol. 11, no. 4, pp. 701-708, Apr. 2017.

[4] F. Fioranelli, M. Ritchie and H. Griffiths, "Bistatic human micro-Doppler signatures for classification of indoor activities," *2017 IEEE Radar Conf. (RadarConf)*, Seattle, WA, 2017, pp. 0610-0615.

[5] B. Jokanovic, M. Amin, F. Ahmad, and B. Boashash, "Radar fall detection using principal component analysis," *Proc. SPIE, Radar Sensor Technology XX*, Baltimore, MD, 2016, vol. 9829, p. 982919.

[6] G. Li, R. Zhang, M. Ritchie, and H. Griffiths, "Sparsity-Driven Micro-Doppler Feature Extraction for Dynamic Hand Gesture Recognition," *IEEE Transactions on Aerospace and Electronic Systems*, accepted and available online.

[7] Y. Kim and L. Hao, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geos. Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, May. 2009.

[8] L. Yang, G. Chen, and G. Li, "Classification of personnel targets with baggage using dual-band radar," *Remote Sensing*, vol. 9, no. 6, p. 594, Jun. 2017.

[9] F. Fioranelli, M. Ritchie, and H. Griffiths, "Centroid features for classification of armed/unarmed multiple personnel using multistatic human micro-Doppler," *IET Radar, Sonar Navigat.*, vol. 10, no. 9, pp. 1702-1710, Dec. 2016.

[10] P. P. Markopoulos and F. Ahmad, "Indoor human motion classification by L1-norm subspaces of micro-Doppler signatures," *2017 IEEE Radar Conf. (RadarConf)*, Seattle, WA, 2017, pp. 1807-1810.

[11] Y. Kim and T. Moon, "Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 1, pp. 8-12, Jan. 2016.

[12] Y. Kim and B. Toomajian, "Hand gesture recognition using micro-Doppler signatures with convolutional neural network," *IEEE Access*, vol. 4, pp. 7125-7130, Oct. 2016.

[13] T. S. Jordan, "Using convolutional neural networks for human activity classification on micro-Doppler radar spectrograms," *Proc. SPIE, Sensors, and Command, Control, Commun., and Intell. (C3I) Technologies for Homeland Security, Defense, and Law Enforcement Applicat. XV*, 2016, vol. 9825, p.982509.

[14] J. Park, R. J. Javier, T. Moon and Y. Kim, "Micro-Doppler based classification of human aquatic activities via transfer learning of convolutional neural networks," *Sensors*, vol. 16, no. 12, p. 1990, Nov. 2016.

[15] M. S. Seyfioğlu, S. Z. Gürbüz, A. M. Özbayoğlu and M. Yüksel, "Deep learning of micro-Doppler features for aided and unaided gait recognition," *2017 IEEE Radar Conf. (RadarConf)*, Seattle, WA, 2017, pp. 1125-1130.

[16] K. N. Parashar, M. C. Oveneke, M. Rykunov, H. Sahli and A. Bourdoux, "Micro-Doppler feature extraction using convolutional auto-encoders for low latency target classification," *2017 IEEE Radar Conf. (RadarConf)*, Seattle, WA, 2017, pp. 1739-1744.

[17] R. P. Trommel, R. I. A. Harmanny, L. Cifola and J. N. Driessen, "Multi-target human gait classification using deep convolutional neural networks on micro-doppler spectrograms," *2016 European Radar Conf. (EuRAD)*, London, 2016, pp. 81-84.

[18] C. Karabacak, S. Z. Gürbüz, M. B. Guldogan and A. C. Gürbüz, "Multi-aspect angle classification of human radar signatures," *Proc. SPIE, Active and Passive Signatures IV*, 2013, vol. 8734, p.873408.

[19] F. Fioranelli, M. Ritchie and H. Griffiths, "Performance analysis of centroid and SVD features for personnel recognition using multistatic micro-Doppler," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 5, pp. 725-729, May. 2016.

[20] F. Fioranelli, M. Ritchie, S. Z. Gürbüz and H. Griffiths, "Feature diversity for optimized human micro-Doppler classification using multistatic radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 2, pp. 640-654, Apr. 2017.

[21] M. B. Özcan, S. Z. Gürbüz, A. R. Persico, C. Clemente and J. Soraghan, "Performance analysis of co-located and distributed MIMO radar for micro-Doppler classification," *2016 European Radar Conf. (EuRAD)*, London, 2016, pp. 85-88.

[22] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345-1359, Oct. 2010.

[23] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, "Learning and transferring mid-level image representations using convolutional neural networks," *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, 2014, pp. 1717-1724

[24] H. C. Shin et al., "Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1285-1298, May. 2016.

[25] K. Chatfield, K. Simonyan, A. Vedaldi and A. Zisserman, "Return of the devil in the details: delving deep into convolutional nets," *Proc. British Machine Vision Conf. 2014*, Nottingham, 2014, pp. 1-12.

[26] S. Ioffe, C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," *Int. Conf. Machine Learning*, Lille, 2015, pp. 448-456.

[27] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learning Research*, vol. 15, no. 1, pp.1929-1958, Jun. 2014.

[28] A. Vedaldi and K. Lenc, "MatConvNet - Convolutional Neural Networks for MATLAB," *Proc. 23rd ACM Int. Conf. Multimedia*, Brisbane, 2015, pp. 689-692.

[29] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," *3rd Int. Conf. Learning Representations*, San Diego, CA, 2015.