

Eliminate Aspect Angle Variations for Human Activity Recognition using Unsupervised Deep Adaptation Network

Qingchao Chen, Yang Liu, Francesco Fioranelli, Matthew Ritchie, Kevin Chetty

University College London, University of Cambridge, University of Glasgow

{qingchao.chen.13, m.ritchie, k.chetty}@ucl.ac.uk, Francesco.Fioranelli@glasgow.ac.uk, yl504@cam.ac.uk

Abstract—Activity recognition and monitoring using radar micro-Doppler signatures (μ -DS) classification has played a vital part in various security and healthcare applications. In the practical scenario, aspect angle variations of μ -DS increase the data diversity but can be regarded as a distraction factor for activity recognition. The learned feature extractor and classifier will degrade a lot if the test μ -DS is from a different aspect angle from the training dataset. This is because the aspect angle variations between training and test dataset will break the assumption of the classification methods: the training and test data are drawn from the same distribution. This paper aims to eliminate the aspect angle variations by learning aspect angle invariant and meanwhile discriminative features in the bi-static radar geometry using the unlabeled test data. More specifically, we first propose a new problem to train a feature extractor using certain aspect angles but generalizes well for other aspect angles in the test stage. Next, we propose two adaptation networks termed as MMD-DAN and JS-DAN, utilizing two widely used distribution divergence measurements. Finally, we evaluate our experimental setting and methods using experimental data.

I. INTRODUCTION

Detecting human targets and their activities relies on the micro-Doppler signature (μ -DS) analysis of the frequency modulations of the backscattered signal. As the additional frequency modulations induced by rotating and vibrating parts of objects, e.g. wheels of trucks, limbs movement of human targets [1–3], μ -DS and the classification methods have been applied in plenty of fields. It has been shown that movement of different human targets can be distinguished, as well as differences between men and women, people and animals [4, 5]. In addition, μ -DS have been used to distinguish wind turbine blades and the blades of aircraft rotors [6]. It has also been demonstrated how μ -DS of different human target movements can help increase the situational awareness of the ambient assistant living in the healthcare context [7–13].

In fact, μ -DSs and their applications have been investigated over a number of years under this regime. The developments include but are not limited to: novel time-frequency analysis methods, feature extraction methods and classifier designs [11, 12, 14–20]. In spite of these novel designs, very few of them consider aspect angle variations in μ -DS and eliminating them to further boost the activity recognition performance. Previous works by Fioranelli et al. [21–23], Chen et al. [19], and Patel et al. [20] considered to adopt the multi-static radar to learn useful features for motion recognition, considering

the dynamic effects of the aspect angle variation for μ -DS classification. However, they have not consider to boost the activity recognition performance by eliminating the aspect angle variation and learn a aspect angle invariant features for bi-static radar geometries.

μ -DS classification methods make assumptions about the model and task. The most widely adopted one is that the training and test data, features or joint feature and label representations are drawn from the same distribution. Unfortunately the aspect angle variations between training and test data may break this assumption in practical scenarios, for example, if the training data are from aspect angle 1, however all the test samples are from aspect angle 2 and 3, the distributions of training and test data diverge and therefore the classification performance will degrade due to the “unseen” aspect angle variations in the training stage.

To tackle the distribution discrepancy or divergence caused by the “unseen” aspect angle variations in bi-static radar geometry, we first propose a new problem for human activity recognition to extract not only the discriminative features for activity recognition but they have to be aspect angle invariant as well. Then we also propose two deep adaptation networks to reduce the distribution divergence by leveraging the Maximum Mean Discrepancy (MMD) and the Jenson-Shannon(JS) divergence measurements in the network design. More specifically, the contributions of this paper are summarized as follows:

- 1) we formulate a new problem for human activity recognition to eliminate the aspect angle variations.
- 2) we handle this problem utilizing the deep adaptation network (DAN) and apply two widely-used one, termed as MMD-DAN and JS-DAN, leveraging MMD and JS divergence respectively.
- 3) we design various experiments handling aspect angle variations between training and test dataset in the context of recognizing armed or unarmed walking human targets. In addition, we compare the results of MMD-DAN and JS-DAN with the baseline network.

The paper is organized as follows: first we formulate the new problem to eliminate the aspect angle variation in the context of domain adaptation theory in section II. Next, we introduce the proposed two deep adaptation network (DAN) to eliminate the aspect angle variations in section III. Section

IV and V focuses on detailed experiments and results. Finally we conclude the paper in section VI.

II. PROBLEM FORMULATION

μ -DS classification methods make assumptions about the model and task. The most widely adopted one is that features or joint feature and label representations in training and test datasets are drawn from the same distribution. Unfortunately due to the unpredictable aspect angle variations, this assumption rarely holds in practical scenarios.

In this paper, we propose a new problem for human activity recognition: **to train a feature extractor and classifier using certain aspect angles but generalizes well for other aspect angles in the test stage**. Note that we may access the test data samples in this task but not the labels since in practical scenarios, we cannot know the aspect angle or the activity of the target in test stage. To address this problem, we propose the relevant task in the following: **to learn aspect angle invariant features trained using μ -DS from fixed aspect angles but generalizes well to test samples from other aspect angles by eliminating the aspect angle variations**.

More specifically, the task is formalized as follows: the training μ -DS data x^s and label y^s are sampled from distribution P^s on fixed aspect angles (e.g. we only sample training μ -DS from **aspect angle 1** only) and termed as the source data or source domain $D_s = \{(x^s, y^s)\}$; while the test data x^t and label y^t are sampled from the other distribution P^t on different aspect angles (e.g. **aspect angle 2 or/and 3**) and termed as the target data or target domain $D_t = \{(x^t, y^t)\}$. Note that $P^t \neq P^s$ and in the training stage, the test label y^t is not available. In this paper, we further assume that there is single activity recognition task and therefore both domains share the same classification categories, where $y^s \in [1, 2, \dots, N_c]$ and $y^t \in [1, 2, \dots, N_c]$. Given this scenario, the task is equivalent to **learn aspect angle invariant feature representations by two ConvNets T_s and T_t so that their feature distribution discrepancies are reduced, where $P^s(T_s(x^s)) \approx P^s(T_t(x^t))$** .

III. UNSUPERVISED DEEP ADAPTATION NETWORK (DAN) FOR μ -DS CLASSIFICATION

In this section, we introduce two deep adaptation networks(DAN), including MMD-DAN and JS-DAN adopting Maximum Mean Discrepancy and the Jensen-Shannon divergence respectively for learning the aspect-angle-invariant features.

A. Network Architecture

In general, architecture of DAN for μ -DS classification is composed of three modules, including feature extractor network modules T_s and T_t , classifier network C and the distribution matching module D . Figure 1 describes the basic architecture of DAN, where

- T_s and T_t aim to extract features from two domains(μ -DS from two aspect angles) as $T_s(x^s)$ and $T_t(x^t)$;

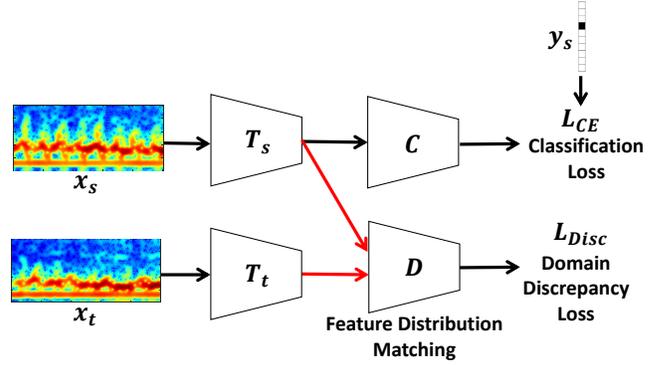


Fig. 1. Architecture of DAN.

- The network/module D then measures the domain distribution discrepancy as $D(T_s(x^s), T_t(x^t))$; The network D and T_t are updated by minimizing the domain discrepancy loss L_{Disc} as shown in Eq.(1).
- The classifier network C takes the source domain feature $T_s(x^s)$ as input, transforms it to predicted label vector $C(T_s(x^s))$. The networks T_s and C are updated by minimizing the cross-entropy (CE) loss L_{CE} in Eq.(2), leveraging the ground-truth source-domain label y^s .

$$\min_{T_t, D} L_{Disc}(D(T_s(x^s), T_t(x^t))) \quad (1)$$

$$\min_{T_s, C} L_{CE}(C(T_s(x^s)), y^s) \quad (2)$$

Note that the MMD and JS based divergence measurements and the relevant L_{Disc} are introduced in the following two sections.

B. MMD-DAN

This section introduces the widely used divergence measurements in adaptation tasks termed as the Maximum Mean Discrepancy (MMD). Let us first assume that N_s and N_t source and target domain samples are drawn from the distributions $P^s(T_s(x^s))$ and $P^t(T_t(x^t))$ respectively.

The MMD is usually considered as the non-parametric estimate of the distance between two distributions, by projecting the input into the Reproducing Kernel Hilbert Space (RKHS) [24]. The projection function can be more formally defined as $\phi(\cdot) : \mathbb{R}^d \rightarrow \mathcal{H}$ with the \mathcal{H} a RKHS and the MMD loss L_{MMD} is formalized in the Eq.(??) where the distribution discrepancy is measured via the difference of the feature sample means in the RKHS. This measurement can be efficiently calculated using the inner product of mapped feature $\phi(T_s(x_i^s))$ and $\phi(T_t(x_j^t))$ via the kernel function. The MMD loss L_{MMD} will only be close to zero if the two distributions under estimate are very similar $P^s(T_s(x^s)) \approx P^t(T_t(x^t))$, as proved in [24]. We may regard L_{MMD} as the discrepancy loss L_{Disc} as shown in Eq.(3).

$$\begin{aligned}
L_{Disc}(x^s, x^t) &= L_{MMD}(x^s, x^t) = D(T_s(x_i^s), T_t(x_j^t)) \\
&= \left\| \frac{1}{N_s} \sum_i \phi(T_s(x_i^s)) - \frac{1}{N_t} \sum_j \phi(T_t(x_j^t)) \right\|_H^2 \quad (3)
\end{aligned}$$

C. JS-DAN

JS divergence may be the first measurement implemented based on the adversarial training strategy. This has been widely used in the well-known Generative Adversarial Network (GAN) [25] and relevant applications. We formulate the JS divergence $JS(P^s || P^t)$ in Eq.(4) and its implementation using adversarial training shown in Eq.(5). Note that KL means the Kullback-Leibler (KL) divergence. We may regard the distribution discrepancy loss L_{Disc} as the JS divergence loss L_{JS} which can be reformulated as the following optimization problem in Eq.(5) where the distribution matching network D tries to measure the JS discrepancy by maximizing the loss L_{JS} .

$$\begin{aligned}
L_{Disc}(x^s, x^t) &= L_{JS}(x^s, x^t) \\
&= JS(P^s || P^t) = KL(P^s || \frac{P^s + P^t}{2}) + KL(P^t || \frac{P^s + P^t}{2}) \quad (4)
\end{aligned}$$

$$\begin{aligned}
L_{Disc}(x^s, x^t) &= L_{JS}(x^s, x^t) \\
&= \max_D \mathbb{E}_{T_s(x^s) \sim P^s} [\log(D(T_s(x^s)))] \\
&\quad + \mathbb{E}_{T_t(x^t) \sim P^t} [\log(1 - D(T_t(x^t)))] \quad (5)
\end{aligned}$$

D. Optimization

To sum up, the overall optimization is formulated as the following Eq.(6). More specifically, MMD-DAN is optimized in Eq.(7) while JS-DAN optimization function is shown in Eq.(8). Intuitively, both MMD-DAN and JS-DAN aim to learn discriminative and aspect angle invariant features by minimizing the CE loss and the distribution discrepancy. However, they use different formulations for the distribution discrepancy loss L_{Disc} , where MMD-DAN utilizes the Kernel function in module D while JS-DAN utilizes the adversarial training strategy and the module D functions like a domain classifier.

$$\min_{T_s, T_t, C} L_{CE} + \lambda_{Disc} L_{Disc} \quad (6)$$

$$\min_{T_s, T_t, C} L_{CE} + \lambda_{Disc} \left\| \frac{1}{N_s} \sum_i \phi(T_s(x_i^s)) - \frac{1}{N_t} \sum_j \phi(T_t(x_j^t)) \right\|_H^2 \quad (7)$$

$$\begin{aligned}
\min_{T_s, T_t, C} \max_D L_{CE} + \lambda_{Disc} \mathbb{E}_{T_s(x^s) \sim P^s} [\log(D(T_s(x^s)))] \\
+ \lambda_{Disc} \mathbb{E}_{T_t(x^t) \sim P^t} [\log(1 - D(T_t(x^t)))] \quad (8)
\end{aligned}$$

IV. EXPERIMENT

A. Radar System and Experiment Scenario

In the following Figure 2, the data are collected when targets walk from three different angles roughly around 30

degree, 0 degree and -30 degree, (denoted as Ang 1,2,3). The distance between the target position and the RX node is 70m and the TX node is 40 meter away from the RX node. There are two movements in the experiment, which are walking with armed and unarmed weapon (replaced with a metal stick holding or not holding). There are three people involved into the experiments, their height are 1.87m, 1.7m and 1.75m respectively. For each walking, the recorded time is 5 second and in total 90 data samples (5 second recording) are collected.

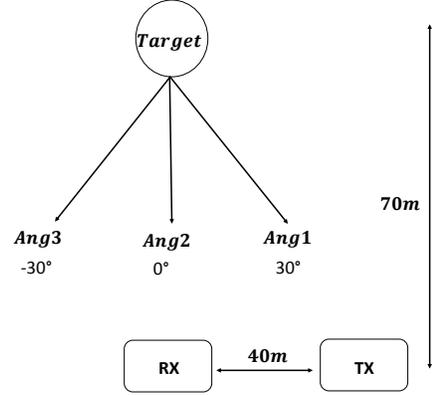


Fig. 2. Experiment scenario in the field using NetRad.

In this paragraph, we introduce the NetRad radar systems, the experimental scenarios in the trial and finally the processing methods and details.

The radar system used to collect the data presented in this paper is the bi-static system from NetRAD [23, 26], which has been developed over the past years at University College London and the basic architecture can be found in [26]. The system is a coherent pulsed radar and operates at 2.4 GHz. The data shown in this work were collected using the following RF parameters: 0.6 μ s pulse duration, 45 MHz bandwidth, linear up-chirp modulation, and 5 kHz PRF to include the whole human μ -DS within the unambiguous Doppler region. Five seconds of data were recorded for each measurement in order to collect a multiple periods of the average human walking gait, which is on average approximately 0.6 seconds. The transmitted power of the radar is approximately 200 mW. The antennas have 24 dBi gain and are operated with vertical polarization to effectively interact with human subjects, as the human body shape is such that the vertical dimension is more significant than the horizontal dimension. This is expected to increase the SNR of the targets' echoes in comparison with horizontal polarization

B. μ -DS Data Processing and Augmentation for DAN

First the matched filter processing is used by cross-correlating the reference and received echo signals and the Short Time Fourier Transform (STFT) is adopted to obtain the spectrogram. The processing parameters is summarized in

TABLE I
PROCESSING PARAMETERS INCLUDING THE ONES TO COLLECT AND GENERATE THE μ -DS, TRAIN THE DCNN AND AUGMENTATION.

Recording Time	5s
Overlapping Ratio	0.9
FFT Integration Time	0.3s
Augmentation Time Width	0.15s
SGD Momentum	0.9
Base Learning Rate	0.001 (FC), 0.0005 (Conv)

Table I, where the overlapping ratio is 0.9 and the integration time of FFT is 0.3 seconds. Each μ -DS sample is recorded for 5 seconds and the stride for cropping the μ -DS samples in the data augmentation operation is chosen as 0.15 seconds. In addition, to increase the challenge of testing, we are also cropping the testing data into different dwell time but the stride is chosen as 0.3 seconds and the cropping starting point is chosen randomly. We argue that this test scheme is a more realistic scenario, where we cannot guarantee where the real-time test data starts, as the radar may have been performing other tasks prior to extracting the μ -DS of a specific target at a specific time.

The basic networks T_s and T_t and data operations are implemented using the Tensorflow software. The network structure of T_s and T_t consists of two Convolutional (Conv) layers, composed of 64 and 128 kernel filters with the size of 32×32 and 11×11 respectively. Following the Conv layers are the Fully Connected (FC) layers which usually transform the local activation maps of Conv layers to the label embedding. We adopt a two-layer architecture with output activation number of 512 and 128 respectively. The network C projects the feature dimension to the representation of semantic categories (two classes in this application). The network D is a single FC layer, which projects the feature to classify the two domains with the output dimension of 2. The conventional SGD method is used for optimizing the parameters, with the momentum 0.9. The initialized learning rate for FC layers and Conv layers are chosen as 0.001 and 0.0005 respectively. The decay policy for the learning rate is the inverse decay and the decayed learning rate denoted as lr_{decay} is following the Eq.(9), where lr_{base} is initialized base learning rate. The batch size is chosen as 50 and training and test samples are shuffled by the Tensorflow FIFO-Queue operation. The regularization weight for the Conv and FC layers are chosen as 0.005.

$$lr_{decay} = lr_{base} \times (1 + 0.001 \times epoch)^{-0.75} \quad (9)$$

Due to the limited number of μ -DS, we design an augmentation method by cropping the original training data into smaller patches, as shown in Figure 4. Due to the nature of the time series of the μ -DS data, we generated more training samples by shifting them in the time domain by different stride sizes. As our training μ -DS data is $\mathbf{x} \in \mathbb{R}^{w_{input} \times h_{input}}$, the augmented training samples can be represented as the cropped data along the time axis via different strides, as shown in Eq.(??), where w_{win} , h_{win} are the window sizes

in two dimensions. This operation, if stride sizes small enough are chosen, will increase the number of training samples, give additional data diversity, and improve the robustness of the model as data generated under various conditions will be used for training. In practice, this time shifting simulates misalignment in time and small Doppler offsets for the training data and these two situations can practically happen in realistic uncontrolled scenarios.

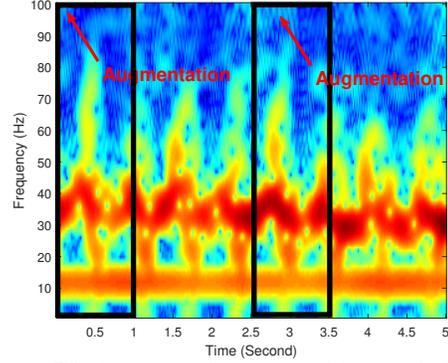


Fig. 3. Raw μ -DS of a target walking unarmored from angle 1, using receiver node 1 in Figure 2; two black boxes indicate the augmented μ -DS, with window size of 1s and window height of 100Hz.

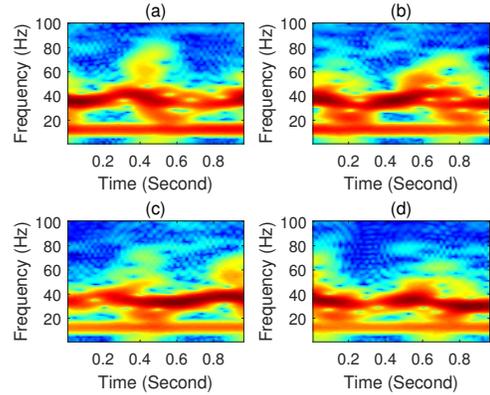


Fig. 4. Augmentation results: (a), (b), (c) and (d) are four augmented data examples generated from the 5-second μ -DS in Figure 3. The augmentation method and parameters are illustrated in Section ??.

More specifically, as shown in Figure 3, an example of a target walking unarmored from angle 1, received by node 1 is illustrated. Here, in the 5-second μ -DS, red and green rectangular boxes indicate two augmented data samples in the training stage. The example shown in Figure 3 uses w_{win} as 1.5 second, while the h_{win} chosen as 100 Hz. It seems obvious that, the augmented data samples can be generated by selecting very small stride of the moving window, which will also simulate small misalignment in realistic data.

C. Experiment Settings to Eliminate Aspect Angle Variations

We follow the experiment setting in Section IV-A and the summary of the utilized μ -DS dataset is in Table II. In this

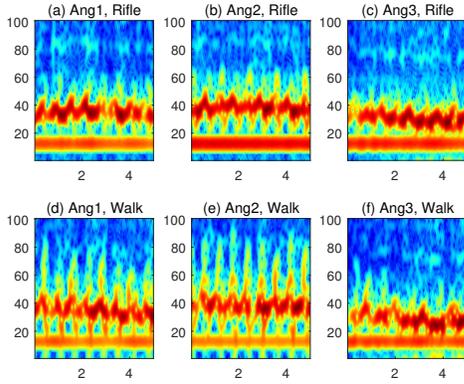


Fig. 5. Raw μ -DS of walking and walking with rifle from three angles. All x-axis represents time with unit of second while all y-axis represents frequency with unit of Hz. The scale is in dB.

section, we mainly evaluate the effect of the MMD-DAN and JS-DAN using different mono-static and bi-static μ -DS and specifically, we design 6 adaptation tasks listed in Table III. In these experiments, our interest factor is the motion while the distraction factors is the aspect angle. The designed six experiments mainly investigate the aspect angle variations. We choose to set up the experiments in a ‘‘Factor Control’’ manner where we use the data from *target 1* only if we focus on eliminating the variations caused by *aspect angles*.

TABLE II
EXPERIMENT DETAILS FOR ADAPTING TWO FACTORS OF VARIATIONS INCLUDING ASPECT ANGLE AND TARGET PERSONNEL.

Bi-Static μ -DS (5-second) Number	90
Train Percentage	20%
Dwell Time	1s
Augmented Train μ -DS Number	918
Augmented Test μ -DS Number	648
Aspect Angles (Degree)	-30, 0, 30

TABLE III
EXPERIMENTS TO EVALUATE THE FACTORS OF VARIATION; THE MAIN INTEREST FACTOR IS THE MOTION RECOGNITION WHILE THE DISTRACTION FACTOR IS THE ASPECT ANGLE. WE CONDUCT THE 6 EXPERIMENTS USING μ -DS FROM TARGET 1 ONLY BUT THREE ASPECT ANGLES SO THAT THE VARIATION CAUSED BY TARGET IDENTITIES IS CONTROLLED.

Experiment Number	Variation Factor	Source Domain (Training Set)	Target Domain (Testing Set)
Exp1	Aspect Angle	Angle 1,2	Angle 3
Exp2	Aspect Angle	Angle 1,3	Angle 2
Exp3	Aspect Angle	Angle 2,3	Angle 1
Exp4	Aspect Angle	Angle 1	Angle 2,3
Exp5	Aspect Angle	Angle 2	Angle 1,3
Exp6	Aspect Angle	Angle 3	Angle 1,2

V. RESULT AND ANALYSIS

A. μ -DS Results

In Fig.5, the Doppler frequencies related to the bulk movement from angle 1 and 2 are centred at around 42Hz, while the

TABLE IV
RECOGNITION RATES (%) BY ELIMINATING THE ASPECT ANGLE VARIATIONS; AVG REFERS TO AVERAGE RESULTS OF THE 6 EXPERIMENTS USING A SPECIFIC METHOD. BASE-NET REFERS TO RESULT OF BASE NETWORK IN FIG.1 WITHOUT THE NETWORK *D* AND THE DISTRIBUTION MATCHING BRANCH.

	Exp1	Exp2	Exp3	Exp4	Exp5	Exp6	AVG
Base-Net	71.1	100.0	100.0	88.9	78.9	93.9	88.8
MMD	74.4	100.0	100.0	92.2	79.4	99.4	90.9
JS	71.1	100.0	100.0	90.4	81.1	98.3	90.2

one from angle 3 is around 34Hz. This is due to the relatively larger aspect angle for angle 3 than angle 1 and 2. Either for armed or unarmed gaits, in general, frequency due to arms movement from angle 1 is smaller than angle 2, while the one from angle 3 is much smaller than the angle 1. These can all be explained by the different Doppler aspect angles in the bi-static radar geometry and the velocity components of the bi-static bisector.

It can be clearly seen that the unarmed walking gait from angle 1 and 2 are clearly distinguished from the armed ones by the signatures from 20Hz to 30Hz caused by the swinging arms. From angle 3, as shown in (c) and (f), the difference between unarmed and armed one is not obvious, but some vague differences in the μ -DS from 25Hz to 35Hz still exist.

B. Results of Eliminating Aspect Angle Variations

For base-net results, it can be observed that it achieves 100.0% in Exp2 and 3, however much lower for other experiments. We hypothesize the reason to be the similarities between aspect angle 1 and 2 and μ -DS from angle 1 and 2 are separated in training and test set respectively in Exp2 and 3. However, for other scenarios, where aspect angle 1 and 2 are both in either the test or training set, the performance degrades to the lowest 71.1%, which demonstrates that the aspect angle variations may degrade the activity classification performance.

However in Table IV, through the distribution matching module by MMD and JS, MMD-DAN and JS-DAN outperform the base-net results for all experiments except Exp 2 and 3 (since the base-net result has been 100.0%). For Exp1 and Exp4-6, the best-performing DAN results outperform the baseline by 3.3%, 3.3%, 2.2% and 5.5% respectively. These results prove the effectiveness of MMD-DAN and JS-DAN to eliminate the aspect angle variations. For comparison between MMD-DAN and JS-DAN, MMD-DAN achieves better results for three experiments however, JS-DAN outperforms MMD-DAN in Exp5 only.

VI. CONCLUSION

This paper first proposes a new problem for human activity recognition using μ -DS classification: to train discriminative and meanwhile aspect angle invariant features in the bi-static radar geometry. Then we proposed two DANs termed as MMD-DAN and JS-DAN for eliminating the aspect angle variations, leveraging the MMD and JS divergence. Through

experimental results, MMD-DAN achieves better and more robust results than JS-DAN in almost all scenarios and different level of aspect angle variations.

REFERENCES

- [1] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Micro-doppler effect in radar: phenomenon, model, and simulation study," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2–21, 2006.
- [2] V. C. Chen, *The micro-Doppler effect in radar*. Artech House, 2011.
- [3] V. C. Chen, W. J. Miceli, and D. Tahmoush, *Radar micro-Doppler signatures: processing and applications*. The Institution of Engineering and Technology, 2014.
- [4] Y. Kim, S. Ha, and J. Kwon, "Human detection using doppler radar based on physical characteristics of targets," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 2, pp. 289–293, 2015.
- [5] R. M. Narayanan and M. Zenaldin, "Radar micro-doppler signatures of various human activities," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1205–1215, 2015.
- [6] F. Fioranelli, M. Ritchie, A. Balleri, and H. Griffiths, "Practical investigation of multiband mono-and bistatic radar signatures of wind turbines," *IET Radar, Sonar & Navigation*, vol. 11, no. 6, pp. 909–921, 2017.
- [7] M. S. Seyfioglu and S. Z. Gurbuz, "Deep neural network initialization methods for micro-doppler classification with low training sample support," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 12, pp. 2462–2466, 2017.
- [8] B. Jokanovic, M. Amin, and B. Erol, "Multiple joint-variable domains recognition of human motion," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0948–0952.
- [9] B. Jokanovic, M. G. Amin, and F. Ahmad, "Effect of data representations on deep learning in fall detection," in *Sensor Array and Multichannel Signal Processing Workshop (SAM), 2016 IEEE*. IEEE, 2016, pp. 1–5.
- [10] B. Jokanovic, M. Amin, and F. Ahmad, "Radar fall motion detection using deep learning," in *Radar Conference (RadarConf), 2016 IEEE*. IEEE, 2016, pp. 1–6.
- [11] Q. Chen, M. Ritchie, Y. Liu, K. Chetty, and K. Woodbridge, "Joint fall and aspect angle recognition using fine-grained micro-doppler classification," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0912–0916.
- [12] Q. Chen, B. Tan, K. Chetty, and K. Woodbridge, "Activity recognition based on micro-doppler signature with in-home wi-fi," in *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*. IEEE, 2016, pp. 1–6.
- [13] F. Fioranelli, M. Ritchie, and H. Griffiths, "Bistatic human micro-doppler signatures for classification of indoor activities," in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0610–0615.
- [14] 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*, 2017.
- [15] D. P. Fairchild and R. M. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 8, no. 5, pp. 425–434, 2014.
- [16] A. Brewster and A. Balleri, "Extraction and analysis of micro-doppler signatures by the empirical mode decomposition," in *Radar Conference (RadarCon), 2015 IEEE*. IEEE, 2015, pp. 0947–0951.
- [17] S. Z. Gurbuz, B. Erol, B. Cagliyan, and B. Tekeli, "Operational assessment and adaptive selection of micro-doppler features," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1196–1204, 2015.
- [18] M. Ritchie, M. Ash, Q. Chen, and K. Chetty, "Through wall radar classification of human micro-doppler using singular value decomposition analysis," *Sensors*, vol. 16, no. 9, p. 1401, 2016.
- [19] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, "Personnel recognition and gait classification based on multistatic micro-doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, 2018.
- [20] J. S. Patel, F. Fioranelli, M. Ritchie, and H. Griffiths, "Multistatic radar classification of armed vs unarmed personnel using neural networks," *Evolving Systems*, pp. 1–10, 2017.
- [21] F. Fioranelli, M. Ritchie, and H. Griffiths, "Performance analysis of centroid and svd features for personnel recognition using multistatic micro-doppler," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 725–729, 2016.
- [22] F. Fioranelli, M. Ritchie, and H. Griffiths, "Multistatic human micro-doppler classification of armed/unarmed personnel," *IET Radar, Sonar & Navigation*, vol. 9, no. 7, pp. 857–865, 2015.
- [23] F. Fioranelli, M. Ritchie, and H. Griffiths, "Aspect angle dependence and multistatic data fusion for micro-doppler classification of armed/unarmed personnel," *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1231–1239, 2015.
- [24] A. Smola, A. Gretton, L. Song, and B. Scholkopf, "A hilbert space embedding for distributions," in *International Conference on Algorithmic Learning Theory*. Springer, 2007, pp. 13–31.
- [25] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.
- [26] S. R. Doughty, "Development and performance evaluation of a multistatic radar system," Ph.D. dissertation, University of London, 2008.