FEATURE DIVERSITY FOR OPTIMIZED HUMAN MICRO-DOPPLER CLASSIFICATION USING MULTISTATIC RADAR

Francesco Fioranelli[(1)], Matthew Ritchie[(2)], Sevgi Zübeyde Gürbüz[(3)], Hugh Griffiths[(2)]

(1) School of Engineering, University of Glasgow, Glasgow, UK
(2) Department of Electronic and Electrical Engineering, University College London, UK
(3) Department of Electrical-Electronics Engineering, TOBB University, Ankara, Turkey

Abstract

This paper investigates the selection of different combinations of features at different multistatic radar nodes, depending on scenario parameters, such as aspect angle to the target and signal-to-noise ratio, and radar parameters, such as dwell time, polarisation, and frequency band. Two sets of experimental data collected with the multistatic radar system NetRAD are analysed for two separate problems, namely the classification of unarmed vs potentially armed multiple personnel, and the personnel recognition of individuals based on walking gait. The results show that the overall classification accuracy can be significantly improved by taking into account feature diversity at each radar node depending on the environmental parameters and target behaviour, in comparison with the conventional approach of selecting the same features for all nodes.

1. Introduction

The micro-Doppler effect refers to the additional frequency components observed in addition to the main Doppler shift of moving targets, which are caused by rotating or vibrating parts such as the propeller of aircraft, wheels of vehicles, or the torso oscillation and swinging of limbs in the case of human targets [1]. Micro-Doppler has been investigated for a variety of applications including search and rescue, security, law enforcement and defence [2-4], but the extraction of suitable information and features from the micro-Doppler signatures and the best methods to exploit these for classification, recognition, and identification are still current research fields [5].

Human micro-Doppler signatures collected by a monostatic radar have been investigated in several works in the literature over the past years. It has been shown how features extracted from the Short Time Fourier Transform (STFT) of these signatures can be used to classify human targets from animals and vehicles in

a ground surveillance radar context [6,7], to discriminate between different activities performed by people such as walking, running, crawling [8-13], and even to identify specific individuals performing the same activity by exploiting the characteristic walking gait and small movement patterns that each individual exhibits [14-16]. Time-frequency transforms [4] other than STFTs have been also proposed to characterise micro-Doppler signatures, such as the Gabor transform, Wigner-Ville transform, Cohen's class time-frequency distributions [17] or Empirical Mode Decomposition [18, 19], all of which have been shown to be effective in representing minute movements [18].

It is well known that the micro-Doppler signature depends on the cosine of the angle between the trajectory of the target and the radar line-of-sight (aspect angle); hence, the classification performance can be compromised when this angle is close to 90° and the micro-Doppler signature is significantly attenuated [11]. When this angle has smaller values, up to approximately 30°, the micro-Doppler signature is reduced but is still usable for successful feature extraction and classification, as shown in [8]. In this context, bistatic and multistatic radar systems have been suggested as a suitable tool to mitigate the detrimental effect of less favourable aspect angles for micro-Doppler based classification, as different radar nodes could be deployed to have at least one node with a suitable view of the target of interest. Experimental research on multistatic/bistatic human micro-Doppler signatures is rather limited. The work in [20, 21] used simulated data based on the Boulic kinematic model to create a single spectrogram for feature extraction and classification by fusing individual spectrograms from different radar nodes. The same multistatic radar system used in this work was employed to collect experimental micro-Doppler signatures of people running and walking in different directions and to compare them with simulated results. Moreover, by analysing the correlation between different channels, the work demonstrated that multistatic signatures actually provided additional information than corresponding monostatic signatures, and suggested that techniques for automatic target recognition were expected to yield better results by exploiting this additional information [22]. The work in [23] proposed a bistatic radar system with two receivers to infer the oscillation trajectory of mechanical objects (e.g. a pendulum) and the facing direction of human subjects performing more complex movements such as swinging arms or picking up objects.

Our previous work in [24-27] used a multistatic radar system to identify unarmed vs potentially armed personnel, initially in the simplified case of walking on the spot and then for actual realistic walking. Empirical features derived from the spectrograms of the micro-Doppler signatures were investigated, such as bandwidth and period of the signature, and compared with features extracted from the Singular Value Decomposition (SVD) of the spectrogram. The effect of different aspect angles and different approaches of exploiting multistatic information was also investigated, achieving classification accuracy of approximately 90% or higher for the most favourable aspect angles and combinations of features. Physical features, i.e. features that are directly related to the kinematics of the movement analysed and extracted from the spectrograms (Doppler-time plots) were also used in [8-9, 13, 28], in the context of monostatic radar. Other possible features have been proposed in [14, 28], using the Cadence Velocity Diagram (CVD) of the micro-Doppler signatures, or features based on speech processing techniques and transformations such as linear predictive coding (LPC), discrete cosine transform (DCT), and cepstral coefficients [11].

All these different features have been previously proposed to analyse human micro-Doppler signatures, and this leads to the question of how many features are needed to optimize the classification performance for a given problem, how to select them, and what the impact of parameters related to the scenario or the radar system in the feature selection process may be. The work in [11] has shown for instance the impact of parameters such as dwell time, aspect angle, signal-to-noise ratio (SNR), and pulse repetition frequency over a vast set of features for classification of different human activities, whereas our work in [15] has investigated other types of features based on centroid and SVD of the micro-Doppler signatures for personnel recognition and the impact of aspect angle and SNR for different classifiers. In [29] mutual information was used as a metric for computing an importance ranking of features, while in [13] it was shown that mutual-information could be used to select different sets of physical features based upon dwell time, aspect angle, and SNR and improve classification performance for monostatic radar systems.

This work takes a further step forward by exploring the additional degree of freedom provided by multistatic systems in the context of optimal exploitation of feature diversity, where different combinations of features can be selected at each radar node, depending on situational parameters - such as the dwell time, the signal-

to-noise ratio, and the aspect angle - that may vary from node to node. Feature diversity adds a level of complexity to the feature selection problem, but is shown to provide improved classification performance by taking into account the specific operational situation at each radar node. Moreover, this work validates results not through simulations as in [11, 13, 29], but through the analysis of measured, experimental data. The analysis presented in the following sections will be based on data collected in two different field experiments. The former relates to the problem of classifying unarmed vs potentially armed personnel. In contrast to previous work [24-26], in these data there is no single target but two subjects who are simultaneously walking with similar speed and close in space, and one may (or not) be carrying a metallic pole representing a rifle. These data have been briefly analysed in [27], but without considering feature diversity and following the conventional approach of using the same features at each multistatic node. The latter experiment is related to the problem of personnel recognition based on individual walking gait, and analyses data from four different subjects. Twelve different features based on the centroid and the SVD of the micro-Doppler signatures are considered for each radar node, and selected with the aim of optimising the classification performance. A brute-force wrapper approach consisting of testing all the possible combinations at each node and selecting the best one is compared with a filter approach that ranks the possible features based on chosen metrics. The experimental results show that the overall classification performance can be significantly improved by exploiting this feature diversity at different radar nodes, compared with the conventional situation where all the nodes perform the same feature extraction and selection.

The paper is organised as follows. Section 2 describes the radar system and the analysis of the data, focusing on the feature extraction and selection approach and on the four classifiers considered. Section 3 presents the experimental setups for the two sets of data analysed in this paper, and then discusses the results for the two problems of unarmed vs armed classification and for personnel recognition. Section 4 concludes the paper and discusses possible future work.

2. Radar system and data analysis

The multistatic radar system used to gather the data analysed for this paper is the coherent pulsed radar NetRAD, developed at University College London over the past twelve years [30]. NetRAD consists of three separate but identical nodes operating at 2.4 GHz with 45 MHz signal bandwidth. Other relevant RF parameters for these experiments include linear up-chirp modulation with 0.6 μs duration and pulse repetition frequency (PRF) equal to 5 kHz, which provides unambiguous sampling of the whole human micro-Doppler signature. The transmitted power of the radar was approximately +23 dBm. The antennas used had approximately 18° (horizontal) ×19° (vertical) beamwidths and 18 dBi gain.

2.1 Feature extraction

For both experiments the human target signature was extracted from the range-time radar data and processed using Short Time Fourier Transform (STFT) to obtain spectrograms. A 0.3 s Hamming window with 95% overlap was used to calculate the STFTs. The spectrograms were divided into blocks of different durations from 1 s to 5 s in 0.5 s steps to represent different radar dwell times for feature extraction and to investigate the effect of this parameter on the overall classification performance.

Although many features have been proposed in the literature for human micro-Doppler [14, 28], in this work features that could be extracted automatically from the spectrograms are investigated, i.e. features that can be evaluated without any pre-processing steps or the use of empirical thresholds, such as those employed in the extraction of physical features like bandwidth or periodicity [26]. The aim of this paper is not investigating all the possible choices of features, as it always possible to have different handcrafted features, but focusing on a subset of automatically extracted features and how choosing a different set at each multistatic radar node can provide an enhancement in performance. More specifically, features based on the centroid of the micro-Doppler signature and on the bandwidth around this centroid have been shown to provide good classification results for recognition of individuals based on their walking gait [15]. In this case four features, namely the mean and standard deviation of both the Doppler centroid and bandwidth were used as input to the classifiers. These features have also been shown to be useful for classification in other domains [31], where it was shown how two features, namely the mean of the Doppler centroid and

bandwidth, could be potentially used to discriminate between micro-drones hovering and flying while carrying different types of payloads. The Doppler centroid can be considered to be an estimate of the centre of gravity of the micro-Doppler signature, and the Doppler bandwidth calculates the energy extent of the micro-Doppler signature around the centroid, as described in [32], where these parameters were applied to characterise the signatures of wind turbines. Equations (1) and (2) show the calculation of these parameters, where $F(i,j)$ represents the value of the spectrogram at the i^{th} Doppler bin and j^{th} time bin and $f(i)$ is the value of the Doppler frequency at the i^{th} bin.

$$f_c(j) = \frac{\sum_i f(i)F(i,j)}{\sum_i F(i,j)} \tag{1}$$

$$B_c(j) = \sqrt{\frac{\sum_i (f(i)-f_c(j))^2 F(i,j)}{\sum_i F(i,j)}} \tag{2}$$

Four features based on the centroid of the micro-Doppler signatures are considered in this paper, namely:

1. Mean of the Doppler bandwidth
2. Mean of the Doppler centroid
3. Standard deviation of the Doppler bandwidth
4. Standard deviation of the Doppler centroid

Additionally, features based on Singular Value Decomposition (SVD) of the spectrograms have also been used for target recognition based on micro-Doppler signatures. Assuming that each spectrogram is a Doppler-time matrix $\boldsymbol{F}$ with dimensions $d \times t$, the SVD decomposition of this matrix will be as indicated in equation (3), where $\boldsymbol{S}$ is a $d \times t$ diagonal matrix with the singular values of $\boldsymbol{F}$, and $\boldsymbol{V}$ and $\boldsymbol{U}$ (with dimensions $t \times t$ and $d \times d$ respectively) are the matrices containing the right and left singular vectors of $\boldsymbol{F}$. The spectrograms are converted into logarithmic scale and normalised to their maximum value prior to applying the SVD decomposition.

$$\boldsymbol{F} = \boldsymbol{USV}^T \tag{3}$$

The work in [33] has shown how the SVD decomposition of the spectrograms can help reduce the dimensionality of the feature space by mapping the most significant information on the singular vectors related to the largest singular values. In particular, it was highlighted how the first three left and right individual singular vectors provided information on the physical characteristics of small Unmanned Aerial Vehicles (UAVs), such as blade rotation periodicity, velocity, and overall micro-Doppler bandwidth. In a similar way, our previous results in [26] have used the standard deviation of the first right singular vector as a single feature to discriminate between unarmed and potentially armed personnel. A different approach was taken in [15], where it was assumed that the relevant information for classification was not concentrated in individual singular vectors, but spread across multiple vectors in the whole $\boldsymbol{U}$ and $\boldsymbol{V}$ matrices. In that case the sum of the element intensity of the $\boldsymbol{U}$ matrix appeared to be a suitable feature for personnel identification based on the individual walking gait represented in the spectrograms. Based on these previous works, eight additional features are considered in this work, namely:

5. Standard deviation of the first right singular vector
6. Mean of the first right singular vector
7. Standard deviation of the first left singular vector
8. Mean of the first left singular vector
9. Standard deviation of the diagonal of the U matrix
10. Mean of the diagonal of the U matrix
11. Sum of pixels of the matrix U
12. Sum of pixels of the matrix V

A total of 12 features were therefore considered, and their samples were extracted from each spectrogram or portion of spectrogram for a chosen dwell time. Among the many different types of features previously proposed, these features were selected from previous work by the authors as those providing good accuracy for similar classification problems. The aim of this work is investigating the effect of using different combinations of these features at each multistatic radar node (feature diversity), as a function of different parameters, such as the dwell time and the aspect angle, which may vary from node to node. The

experimental results presented in the following sections show that the overall classification performance can be improved exploiting this feature diversity at each node, in comparison with the approach of using the same feature or set of features for all nodes. The number of considered features was limited to twelve to have a reasonable computational burden when testing all possible feature combinations at the three multistatic radar nodes. However, the methodology of feature diversity can be extended to an initial feature set of any size.

2.2 Classifiers

Four different classifiers were used to process the data presented in this paper, namely Naïve Bayes (NB), diagonal-linear variant of discriminant analysis (DL), nearest-neighbour with 3 neighbours (KNN), and binary trees (BT). A detailed description of these classifiers can be found in [34, 35]. The NB classifier assumes that the feature samples of each class are Gaussian distributed and statistically independent, and that the mean μ and variance σ^2 of these distributions can be estimated from the training data, as shown in equation (4), where x_i indicates the training samples for the i^{th} class. Then the posterior probability of each sample under test belonging to each class is calculated, and the sample is assigned to the class showing the highest posterior probability [35].

$$\mu_i = \frac{1}{N}\sum_n^N x_{i,n} \quad \sigma_i^2 = \frac{1}{N}\sum_n^N \left(x_{i,n} - \mu_i\right)^2 \tag{4}$$

In a similar way, the DL classifier also assumes that the feature samples of each class can be modelled by a multivariate Gaussian distribution as in equation (5), and its mean μ_k and covariance matrix Σ_k can be estimated at the initial training phase of the classifier (the diagonal-linear variant will assume a single covariance matrix for all the classes and estimate only mean values for each class). The sample space is then partitioned into different regions where an expected classification cost C is calculated and minimised with respect to each predicted classification as in equation (6), where $\hat{H}$ is the classification posterior probability.

$$P_{(x|k)} = \frac{1}{\sqrt{2\pi|\Sigma_k|}} exp\left(-\frac{1}{2}(x-\mu_k)^T \Sigma_k^{-1}(x-\mu_k)\right) \tag{5}$$

$$\hat{y} = \underset{y=1,\ldots,K}{argmin} \sum_{k=1}^{K} \hat{H}(k|x) C(y|k) \quad (6)$$

The KNN classifier calculates the Euclidean distance between the samples containing training data for the classifier and the test samples, as indicated in equation (7) where x_s is the vector containing training samples of the i^{th} class and x_t is the vector containing samples under test. The 3 smallest distances are selected for each sample under test, and this is classified as belonging to the class that generated the highest number of these distances, in this case at least 2 out of 3.

$$d_i = \sqrt{\sum \left| x_{i,s} - x_t \right|^2} \quad (7)$$

The BT classifier uses a decision tree for classification of samples under test using binary splits from the root node down to a leaf node, which assigns these samples to a certain class. At the training stage, the tree is built by considering all the possible binary splits on all the available feature samples and selecting the best split according to an optimization criterion. This procedure is then recursively repeated on the two child nodes, until the resulting child node is a 'pure' node, with samples belonging just to a single class. The optimization criterion used is the Gini's Diversity Index (GDI), defined as in equation (8), where *i* denotes the i^{th} class and *n* the node, and *p(i)* is the fraction of classes observation belonging to the i^{th} class that reaches that node. Therefore, if a node is pure and contains only observations of one class, its GDI will be equal to 0, otherwise it is generally a positive number.

$$GDI_n = 1 - \sum_i p^2(i) \quad (8)$$

All classifiers were trained using 25% of the available feature samples and tested on the remaining data. This was done for both the unarmed vs armed classification case and for the personnel recognition case. This small set of data for training was used to investigate the performance of the proposed approach when only a very limited amount of data is available for training and testing, which is often the case for experimental data, especially multistatic data given the practical challenges to operate the system and generate data. The classification error was calculated as the ratio of the overall number of misclassification events and the overall number of samples. This training and validation process was repeated 30 times

selecting random samples for the training data in order to remove possible bias and generalise the performance assessment. The average classification error over these 30 repetitions was calculated and the results are shown in terms of accuracy, i.e. 100% minus the average error. The information available from multiple radar nodes was fused using a binary voting procedure. Each classifier is implemented separately with the data from each individual radar node, and the partial decisions are combined to reach the final decision with the majority, equal in this case to two nodes out of three. In case of non-binary decision, such as the personnel recognition problem with 4 subjects, the final decision is taken by the classifier with the highest confidence in case all three partial decisions from the three nodes are all different.

3. Experimental setups and results

3.1 Description of the two experimental setups

The data presented in this paper refer to two different experiments performed in February 2016 and March 2016 respectively, at the UCL sports ground in an open football field to the North of London. The geometry of the first experiment is shown in Fig. 1a. For this experiment the three NetRAD nodes were deployed in a linear baseline with 50 m separation between nodes, with node 1 acting as monostatic transceiver in the middle and node 2 and 3 as bistatic receivers on the sides. Vertical polarisation was used at all nodes in this experiment. The targets were two people walking together at approximately 70 m from the baseline and moving on five different trajectories, with five different aspect angles with respect to the baseline as indicated in Fig. 1a. Two classes of data were collected, the former with both people walking free handed (‘unarmed’ case), and the latter with one of the two people carrying a metallic pole representing a rifle (‘armed’ case). In different recordings a different person out of the two subjects carried the pole to obtain increased variability in the micro-Doppler signatures. The pole was of comparable size to that of a real rifle and held with both hands. Fig. 1b shows an example of a single person carrying the pole representing the rifle during the experiment. For this experiment the duration of each recording was 5 s to collect multiple repetitions of the average human walking gait. The total number of recordings was 180, assuming 3 nodes, 5 aspect angles, 2 classes (armed vs unarmed), and 6 repetitions per class. The two subjects were simultaneously moving on the same trajectory, closely in space, and the classification between the unarmed

case and armed case is expected to be challenging, as both targets are in the same range bin (the range resolution is approximately 3.3 m with 45 MHz bandwidth) and overlapped in Doppler (both people walking with comparable speed in the same direction).

For the second experiment the nodes were deployed as shown in Fig. 2, with node 1 (monostatic transceiver) and node 2 (bistatic receiver) co-located, and node 3 (bistatic receiver) separated by 50 m. In this case the chosen polarisation was vertical at all nodes, apart from node 2 that recorded horizontally polarised data so that the overall database consists of monostatic co-polarised and cross-polarised data as well as bistatic co-polarised data. The subjects acting as targets were located further away from the baseline, at approximately 90 m. Four different people took part to this experiment and walked towards the baseline, with the aim of analysing their micro-Doppler signatures for personnel recognition. The key body parameters of these subjects were 1.87 m, male, average body type for person A, 1.60 m, female, average body type for person B, 1.78 m, male, slim body type for person C, and 1.70 m, male, average body type for person D. The duration of each recording was 10 s for this experiment. The total number of recordings was 120, assuming 4 people, 3 nodes, and 10 repetitions for each subject.

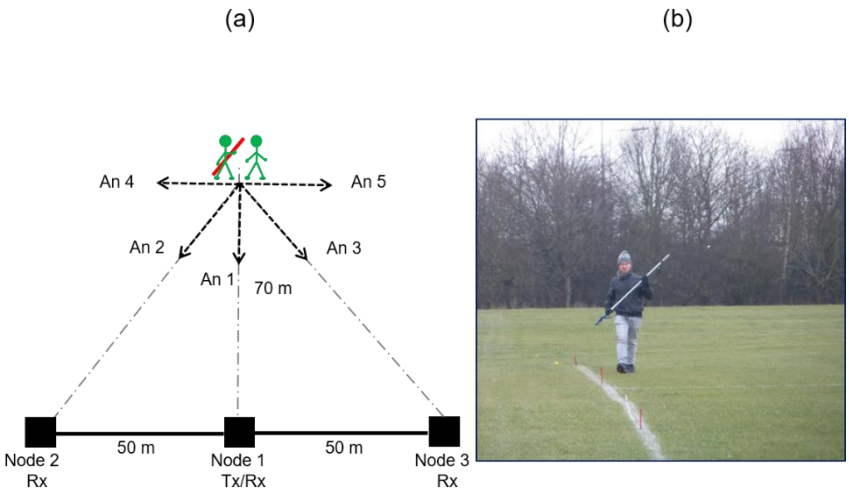

Figure 1 Measurement setup for unarmed vs armed classification experiment (a), and example of person carrying the pole representing the rifle (b)

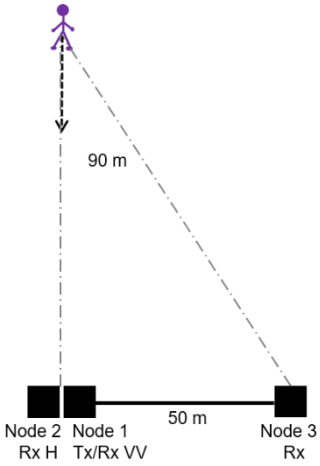

Figure 2 Measurement setup for personnel recognition experiment

3.2 Classification of unarmed vs armed personnel

Fig. 3 shows two examples of spectrograms for the unarmed vs potentially armed classification problem, where two people are walking towards the radar simultaneously and closely in space, and one of them may be carrying the metallic pole representing a rifle. The data used to generate Fig. 3 refer to aspect angle 1 as shown in Fig. 1a, and were recorded at the monostatic transceiver node. The main component of the micro-Doppler signatures in both unarmed and armed cases is at approximately 30 Hz, corresponding to a walking speed of 1.88 m/s, which is reasonable for adults walking at a steady pace. However, the signatures of the two people appear to be overlapped and indistinguishable from the spectrograms, and it should be noted that this was also the case in the range-time domain as the two subjects were closer to each other than the range resolution of the radar.

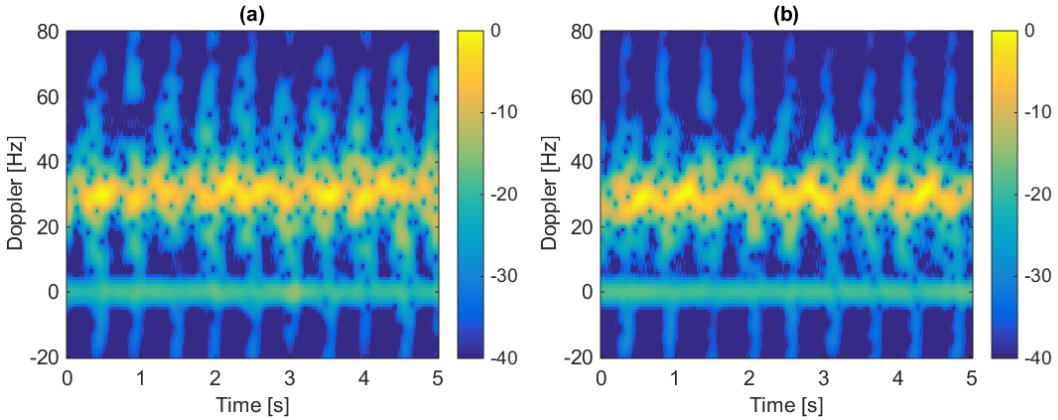

Figure 3 Spectrograms recorded at the monostatic node for two people walking together: both unarmed (a) and one armed and one unarmed (b)

The features described in the previous section were extracted and processed by the four classifiers, with the aim of assessing the effect of exploiting different features at each multistatic radar node. Considering the 12 aforementioned features and assuming to use initially 1 feature per node, there are 1728 (12^3) combinations to test for each classifier following a wrapper approach, i.e. brute force approach of performing all possible tests and selecting the final best result. Fig. 4 shows examples of how the classification accuracy changes depending on the combinations of features used at each multistatic radar node, with the constraint of using a single feature per node. The results from the NB classifier were used in this case, and the red line denotes the average classification. Given a certain dwell time and aspect angle, it can be seen that the accuracy can change significantly, more than 20%, depending on the combination of features used at multiple radar nodes. This shows how the optimal selection of features has an extra level of complexity in multistatic systems, but can deliver improved performance if knowledge of the most suitable combination of features can be inferred or obtained for a certain scenario.

Fig. 5 shows the best classification accuracy obtained for the different aspect angles, classifiers, and dwell times considered in this work. A first observation is that the performance is fairly uniform with different classifiers, and there are no very significant differences between the patterns in the four sub-figures. Aspect angles 1, 2, and 3 appear to provide higher classification accuracy compared with angle 4 and 5. This was expected, as these last two trajectories were parallel to the baseline, hence the spread of the micro-Doppler signature was reduced impacting the feature extraction process. It is interesting to observe the effect of the dwell time, with in general an increase in accuracy with longer dwell times, but this is more relevant for the least favourable aspect angles, e.g. the accuracy shows a significant increase for aspect angle 4 and dwell times longer than 2.5 s, but the pattern remains fairly flat as a function of dwell time for aspect angle 2 and 3 (favourable aspect angles). It is also interesting to notice that the best classification performance is obtained at angle 2 and angle 3, rather than at angle 1, which corresponds to walking straight towards the radar nodes at the baseline. This may be related to the additional information extracted from the signatures collected at different nodes when the individuals were walking towards node 2 or 3 (angle 2 and 3 respectively), in comparison to the symmetric setup of angle 1 with the transceiver node in the middle, but

additional tests are necessary to fully understand and characterise this result. Table 1 summarises the accuracy obtained for the BT classifier as shown in Fig. 5d together with the features used at each node, i.e. for a given aspect angle and dwell time, each group of three numbers indicates which feature out of the list in the previous section was used at node 1, node 2, and node 3, respectively. These features were identified with the wrapper approach of testing all the possible combinations and selecting those yielding the best result. It is interesting to observe how the best selected feature changes for different aspect angles and dwell times. Some features seem to be very recurrent at certain aspect angles but not at others (e.g. feature 2 is almost always selected at node 1 for aspect angle 1 but not used almost at all for other aspect angles), and often – but not always – the best features are different for each multistatic radar node even in the same conditions of aspect angle and dwell time. In Fig. 6 the classification accuracy for three aspect angles as a function of dwell time is reported, with the aim of comparing the optimal accuracy obtained by the brute force wrapper approach of testing all possible combinations of single feature per node, with a possible sub-optimal approach of using the best feature at the monostatic node for all radar nodes (indicated as ‘mono features’ in Fig. 6). Results from the NB classifier were used for this figure. The degradation in classification performance between the two approaches can be seen for all considered aspect angles and dwell times, i.e. forcing all nodes to use the best feature at the monostatic node appears to provide significant reduction in overall accuracy. It is important to consider the added degree of freedom and inherent complexity in exploring this ‘feature diversity’ for classification using multistatic radar systems.

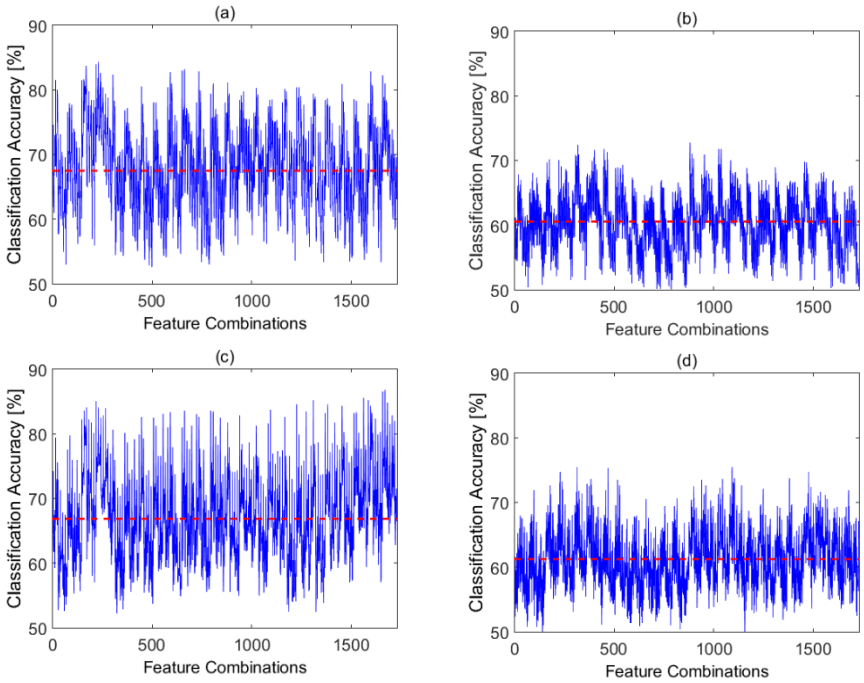

Figure 4 Classification accuracy vs different combinations of single features used at each radar node: (a) angle 1 and dwell time 1 s, (b) angle 4 and dwell time 1 s, (c) angle 1 and dwell time 2.5 s, and (d) angle 4 and dwell time 2.5 s

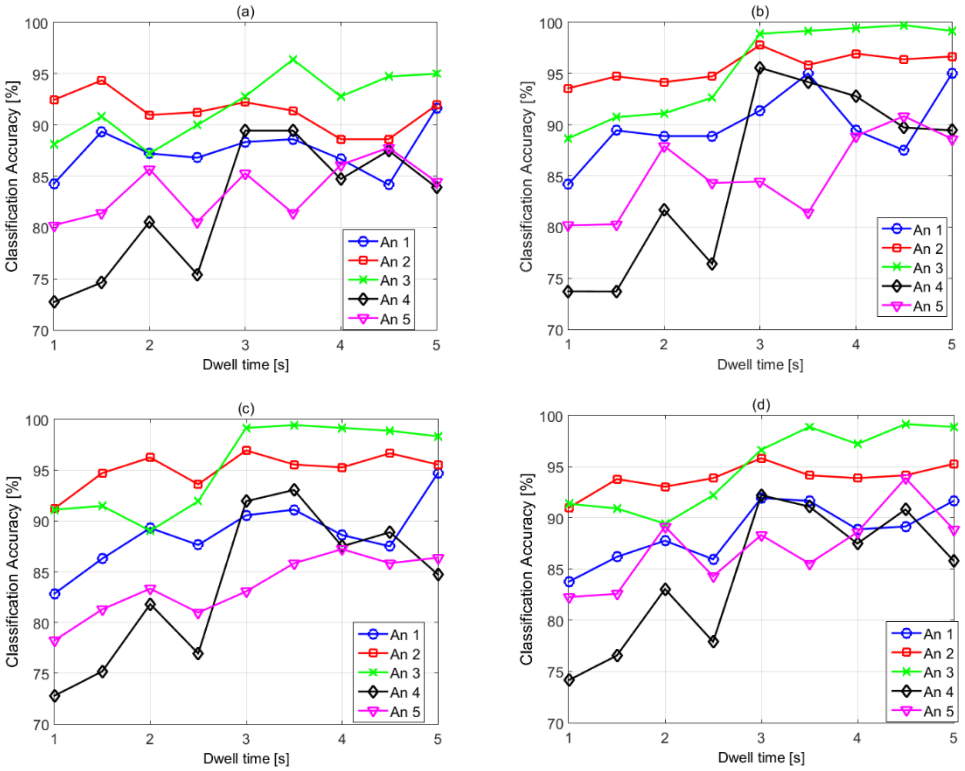

Figure 5 Classification accuracy vs dwell time using the best single feature at each radar node: (a) NB classifier, (b) DL classifier, (c) KNN classifier, and (d) BT classifier

Table 1 Classification accuracy vs dwell time and aspect angle using BT classifier. The single feature used at each radar node is also indicated.

Classification accuracy [%]		1 s	1.5 s	2 s	2.5 s	3 s	3.5 s	4s	4.5 s	5 s
An 1	Accuracy	83.8	86.2	87.8	86.0	91.9	91.7	88.9	89.2	91.7
	Features	2-8-5	2-11-5	2-2-11	2-2-10	2-9-12	2-10-2	10-9-2	2-1-2	2-11-10
An 2	Accuracy	91.0	93.8	93.1	93.9	95.8	94.2	93.9	94.2	95.3
	Features	1-8-2	1-2-2	1-2-2	1-12-2	9-1-2	1-2-2	9-1-2	9-1-2	1-4-2
An 3	Accuracy	91.4	90.9	89.4	92.2	96.7	98.9	97.2	99.2	98.9
	Features	1-3-11	1-11-11	1-2-2	4-11-11	7-2-2	8-2-2	8-3-2	1-1-8	1-11-1
An 4	Accuracy	74.2	76.6	83.1	77.9	92.2	91.1	87.5	90.8	85.8
	Features	4-3-3	10-7-12	5-11-2	10-7-2	8-12-3	7-8-2	7-12-2	2-12-2	13-3-3
An 5	Accuracy	82.3	82.6	89.2	84.3	88.3	85.6	88.6	93.9	88.9
	Features	3-11-2	3-11-2	3-11-2	12-7-2	3-4-4	4-2-3	10-3-2	11-1-3	11-12-2

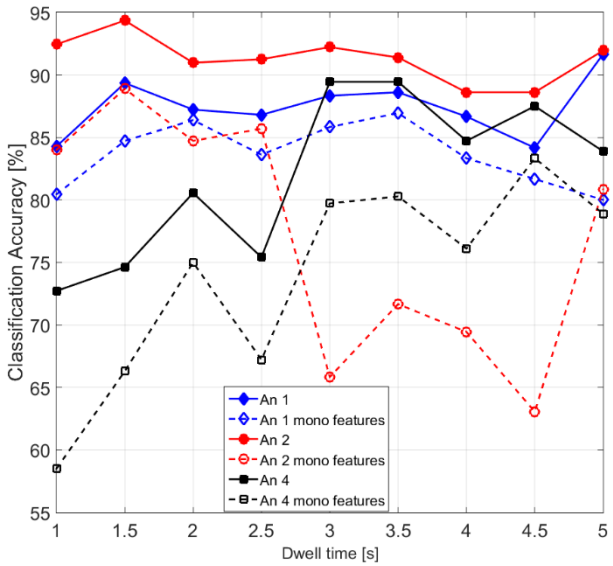

Figure 6 Classification accuracy comparison using best feature combination selected by wrapper approach vs using best combination for monostatic node at all radar nodes (indicated as 'mono features'). NB classifier with single feature per radar node was used.

The analysis has been extended to considering multiple features to be used at each node, which are added with a sequential forward selection (SFS) approach. Initially a wrapper brute-force test of 1728 combinations (i.e. testing all the possible combinations of 12 single features at 3 nodes) identified the best single feature at each node. Then a second feature at each node can be added from the pool of the remaining 11 features, leading eventually to use a pair of features at each node. This implies an additional testing of 1331 combinations (11^3) for each classifier. In a similar way a third feature has been added for each radar node, testing additional 1000 (10^3) combinations per classifier. Fig. 7 presents example of results for two representative aspect angles (angle 1 for favourable Doppler and angle 4 for less favourable Doppler) and

four classifiers, highlighting the differences in performance when using a single feature, pairs of features, and three features at each node. It can be seen that similar trends are observed for the different classifiers considered, and that the effect of dwell time on the accuracy is more evident on aspect angle 4 than on angle 1, i.e. the improvement in accuracy with dwell times longer than 2.5 s is more significant at the less favourable aspect angle.

There is great variability in the results when using or not multiple features, depending on the combinations of the other parameters considered here, i.e. aspect angle, dwell time, and type of classifier. In general, it appears that using pairs of features rather than a single feature at each node can improve the overall accuracy for this particular classification problem and the considered set of features, whereas increasing the number of features used at each node from two to three can in some cases lead to reduced accuracy. This effect of reaching a peak of accuracy with a certain number of features used as input to the classifiers and subsequent plateau or even reduction if more features are used was also observed in other works in the literature [11, 25]. Table 2 shows the actual three features used at each multistatic radar node for the BT classifier case and aspect angle 1 and 4, i.e. for two of the curves shown in Fig. 7d. These were the combinations provided the best classification accuracy as selected through the brute-force wrapper approach of testing the whole number of possible combinations. It is very interesting to observe that these combinations of features vary significantly across the parameters considered, such as aspect angle and dwell time (feature diversity). For example, given an aspect angle, e.g. angle 1, one can see that feature 2 is used pretty much consistently at node 1 for all dwell times, but never used at the bistatic node 2 and only sporadically at the bistatic node 3. There is also a significant variation in features used at the same nodes and aspect angles for different dwell times, as well as significant differences in the features used at the same node and same dwell time, but at different aspect angles. The histograms in Fig. 8 help visualise how different features are used at different nodes for a given aspect angle and classifier, across the considered values of dwell times. It is interesting to observe that some features are selected very often at one node but not at others (e.g. feature 2 used quite often at node 1 and 3 but not at all at node 2), and that some features are not used at all or very rarely. These

results show the importance of being aware and exploit different features at different multistatic radar nodes, depending on the different scenario parameters of the classification problem under test.

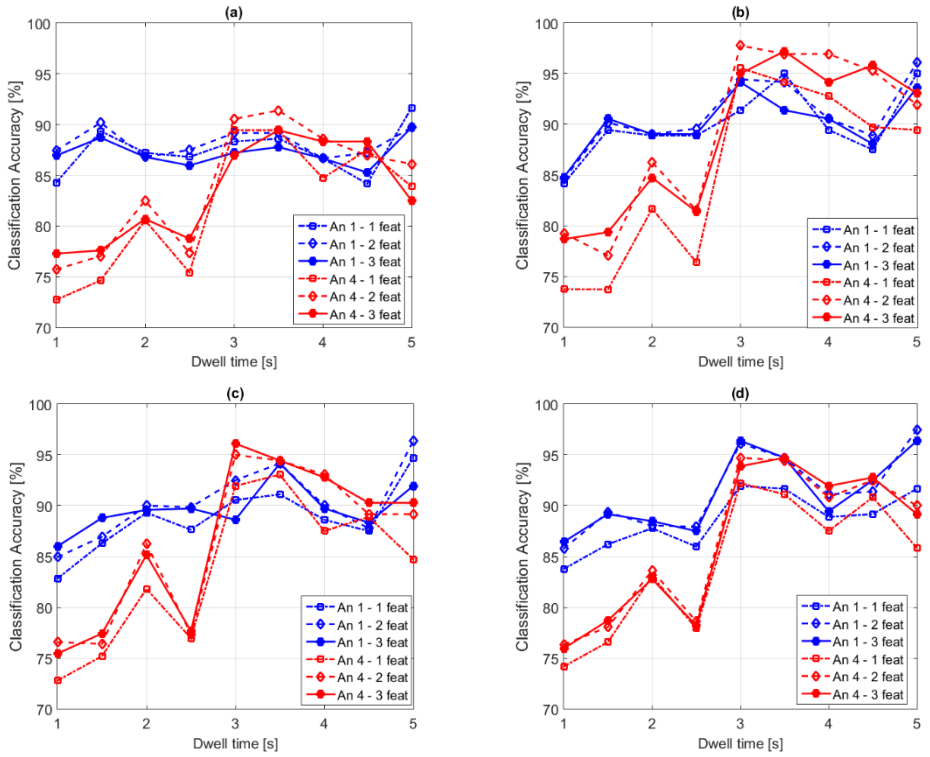

Figure 7 Classification accuracy vs dwell time using the best combination of 1 feature, 2 features, and 3 features per radar node: (a) NB classifier, (b) DL classifier, (c) KNN classifier, and (d) BT classifier

Table 2 Classification accuracy vs dwell time using BT classifier and the best combinations of three features per radar node.

Classification accuracy [%]		**1 s**	**1.5 s**	**2 s**	**2.5 s**	**3 s**	**3.5 s**	**4s**	**4.5 s**	**5 s**
An 1	**Accuracy**	87.1	89.4	88.3	88.2	96.4	95.3	90.0	93.6	96.4
	Features N1	2-11-1	2-9-3	2-7-4	2-4-12	2-7-9	2-3-12	10-1-4	2-3-8	2-3-11
	Features N2	8-1-7	11-1-5	2-8-6	2-11-7	9-10-3	10-4-5	9-12-3	1-4-6	11-12-8
	Features N3	5-2-11	5-11-6	11-2-3	10-11-6	12-5-7	2-10-9	2-10-1	2-6-8	10-6-5
An 4	**Accuracy**	75.5	78.0	82.2	77.8	94.2	95.3	91.4	92.8	89.7
	Features N1	4-6-5	10-3-6	5-10-2	10-3-2	8-2-7	7-12-6	7-8-9	2-9-5	12-2-8
	Features N2	3-2-5	7-5-2	11-4-5	7-8-12	12-8--11	8-5-12	12-8-5	12-5-1	3-6-7
	Features N3	3-10-11	12-3-9	2-3-1	2-1-4	3-11-2	2-3-6	2-11-7	2-8-11	3-8-2

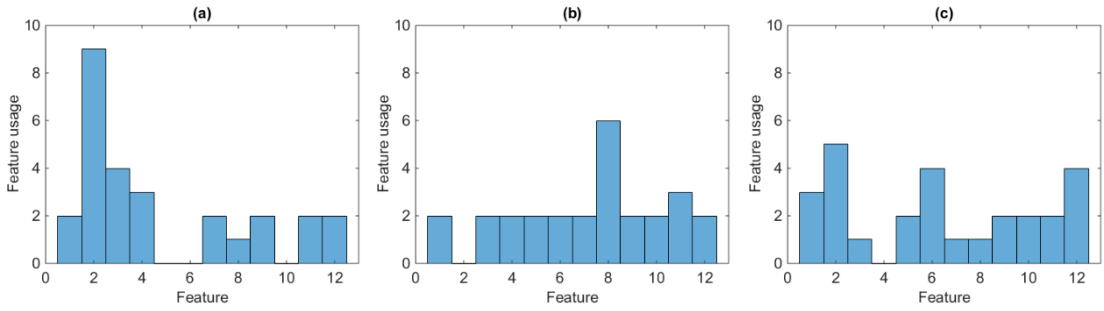

Figure-8 Histograms of features used at each radar node for aspect angle 1, BT classifier, and different dwell times: (a) node 1, (b) node 2, and (c) node 3

The brute-force wrapper approach is computationally very intensive and is tied to the type of classifier used in the tests to evaluate features, so several separability metrics to rank features independently and a priori with respect to classifiers have been proposed in the literature, for instance the T-test and mutual information [11, 13, 36]. In this work two methods of ranking features have been implemented in MATLAB and used separately to rank the samples of the 12 features at the different radar nodes. The first method uses the T-test to compare the mean parameter of two independent groups of data samples, as indicated in equation (9), where x_m and y_m are the means of the groups of samples, s_x and s_y the standard deviations, and N and K are the sample sizes [37].

$$t = \frac{x_m - y_m}{\sqrt{\frac{s_x}{N} + \frac{s_y}{K}}} \tag{9}$$

The second method is based on the relative entropy of the distribution of groups of data samples, which can be related to the concept of mutual information between two random variables and to the Kullback-Leibler divergence [36]. The mutual information and the entropy of discrete random variables X and Y are reported in equations (10) and (11) for completeness, and the details of the mathematical derivation are available in [36, 38]. X and Y are the discrete random variables which can assume N_x (N_y) possible values x_i (y_i) with probability P_X and P_Y, and P_{XY} is the joint probability of the variables X and Y.

$$I(\mathrm{X},\mathrm{Y}) = \sum_i^{N_x} \sum_j^{N_y} P_{XY} \times log_2 \frac{P_{XY}}{P_X P_Y} \tag{10}$$

$$H(X) = -\sum_i^{N_x} P_X log_2 P_X \tag{11}$$

Fig. 9 summarises the classification accuracy for aspect angle 1 and 4 as a function of dwell time, with data generated by the BT classifier. The accuracy obtained using one single feature, pairs of features, and three features at each node is shown, comparing the cases when these features are chosen by the wrapper approach and the ranking approach based on the T-test and entropy criteria. Comparing Fig. 9a to 9c and Fig. 9d to 9f for a given feature selection criterion, one can see that there is an increase in accuracy when adding more features per node, but this is limited in some cases, as already observed with respect to Fig. 7. It is interesting to notice that the performance is very similar in all cases when ranking features with either the T-test or the entropy criterion, but both provide a significant reduction in accuracy compared with the wrapper approach, up to 10-12% in case of the less favourable aspect angle. However, the advantage of filter approaches is that they do not depend on specific classifiers and their implementation. Table 3 shows the accuracy for aspect angle 4 when a single feature is used at each radar node (as in Fig. 9d), and reports the actual features selected by the wrapper and the two considered ranking approaches. The diversity of the features selected with different ranking approaches and wrapper can be seen.

Fig. 10 shows additional results related to two different aspect angles, namely aspect angle 2 (more favourable for Doppler) and aspect angle 5 (less favourable), and two other classifiers, namely KNN and NB. A single feature is used at each radar node. Each sub-figure compares the resulting accuracy by selecting features with a different approach, i.e. the brute-force wrapper, the ranking with the T-test and with the entropy criterion, and the suboptimal approach of forcing all the nodes to use the best feature for the monostatic node as identified by the wrapper method. One can see that the reduction in accuracy can be significant with respect to the optimal wrapper method when using feature selection by ranking, and this is observed for these aspect angles and classifiers in Fig. 10 in addition to the data already observed in Fig. 9. It is interesting to observe that the suboptimal method of forcing all nodes to use the best feature at the monostatic node can provide in some cases better results than the ranking of features either via T-test or via entropy, but not at all the considered dwell times. In any case, the wrapper method provides the best classification accuracy as expected.

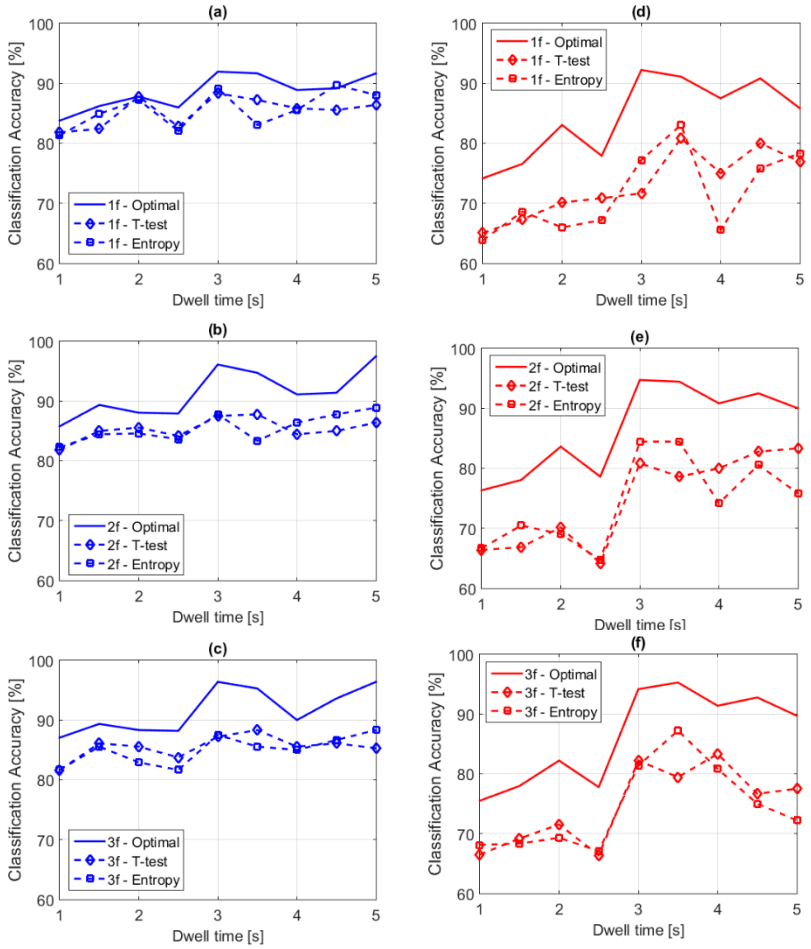

Figure 9 Classification accuracy vs dwell time using BT classifier and different combinations of features per node, selected by wrapper and ranking approaches: (a) 1 feature angle 1, (b) 2 features angle 1, (c) 3 features angle 1, (d) 1 feature angle 4, (e) 2 features angle 4, and (f) 3 features angle 4

Table 3 Classification accuracy vs dwell time using BT classifier and single feature selected at each node using wrapper and ranking approaches. Results related to aspect angle 4.

Classification accuracy [%]		**1 s**	**1.5 s**	**2 s**	**2.5 s**	**3 s**	**3.5 s**	**4s**	**4.5 s**	**5 s**
Optimal	**Accuracy**	74.2	76.6	83.1	77.9	92.2	91.1	87.5	90.8	85.8
	N1	4	10	5	10	8	7	7	2	12
	N2	3	7	11	7	12	8	12	12	3
	N3	3	12	2	2	3	2	2	2	3
T-test	**Accuracy**	65.1	67.3	70.1	70.8	71.7	80.8	75.0	80.0	76.9
	N1	3	2	2	12	2	12	12	2	12
	N2	10	3	11	3	3	11	3	3	3
	N3	3	3	3	3	3	3	2	3	2
Entropy	**Accuracy**	63.8	68.6	66.0	67.2	77.2	83.1	65.6	75.8	78.3
	N1	2	2	5	12	2	12	12	12	12
	N2	2	2	11	3	3	11	6	6	4
	N3	10	10	3	3	2	2	2	2	2

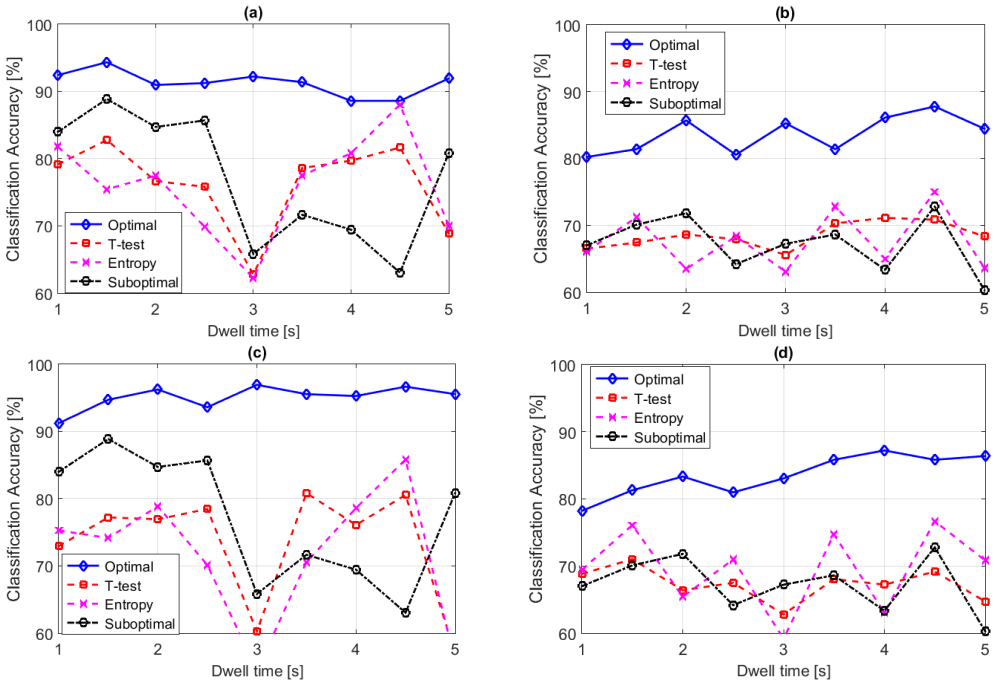

Figure 10 Classification accuracy vs dwell time using a single feature at each radar node selected by wrapper and ranking approaches: (a) angle 2 NB classifier, (b) angle 5 NB classifier, (c) angle 2 KNN classifier, and (d) angle 5 KNN classifier

3.3 Personnel recognition

In this section the classification problem of identifying people from their walking gait is investigated. Fig. 11 shows examples of spectrograms for the four subjects walking towards the radar baseline, as recorded by the monostatic node. In all cases the average speed is between 1.2-2 m/s, corresponding to approximately 20-35 Hz main Doppler shift, which is a realistic value for people walking. Some differences can be empirically seen between these spectrograms, and the analysis in this section shows classification results based on the possible 12 features and 4 classifiers mentioned in the previous section. It is important to notice that node 2 and node 1 were co-located, but operating at different polarisations (node 1 was receiving co-polar V polarisation, and node 2 cross-polar H polarisation). Any difference in feature selection approach between these two nodes is therefore expected to be related to the polarisation diversity rather than to spatial diversity and different locations of bistatic nodes, as in the previous section on unarmed vs armed classification.

Fig. 12 summarises the results in terms of accuracy as function of the dwell time for the four classifiers. Each sub-figure refers to the case of using one single feature, 2 features, 3 features, and 4 features at each radar node. These features have been selected using the brute-force wrapper approach, hence 1728 (12^3)

combinations to test per classifier to select the first feature, then 1331 (11^3) combinations at the second step, 1000 (10^3) at the third step, and 729 (9^3) at the fourth step. One can see the increase in accuracy caused by using additional features from Fig. 12a to Fig. 12d, and how this improvement tends to be less and less significant or become an actual reduction when having more than 3 features per node. This can be seen with more clarity in Fig. 13, which shows in the same figure the results obtained by the BT classifier when a different number of features is used at each radar node. The accuracy pattern appears to be fairly consistent for different classifiers, with the NB and DL classifiers providing the best results. The plots in Fig. 12 show a clearer increasing trend of accuracy as function of the dwell time for the personnel recognition problem compared with similar figures in the previous section on unarmed vs armed personnel classification. These results show an overall classification accuracy above 90% for dwell time longer than approximately 4s and more than 2 features used as input to the classifiers, which is a significant result considering that the personnel recognition problem is in general more challenging than classification of activities.

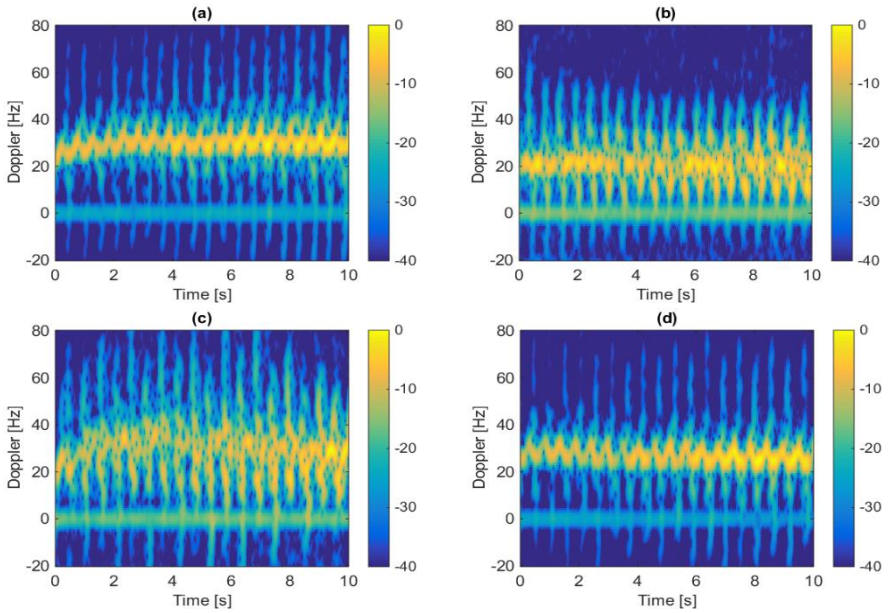

Figure 11 Spectrograms recorded at the monostatic node for different people walking: person 1 (a), person 2 (b), person 3 (c), and person 4 (d)

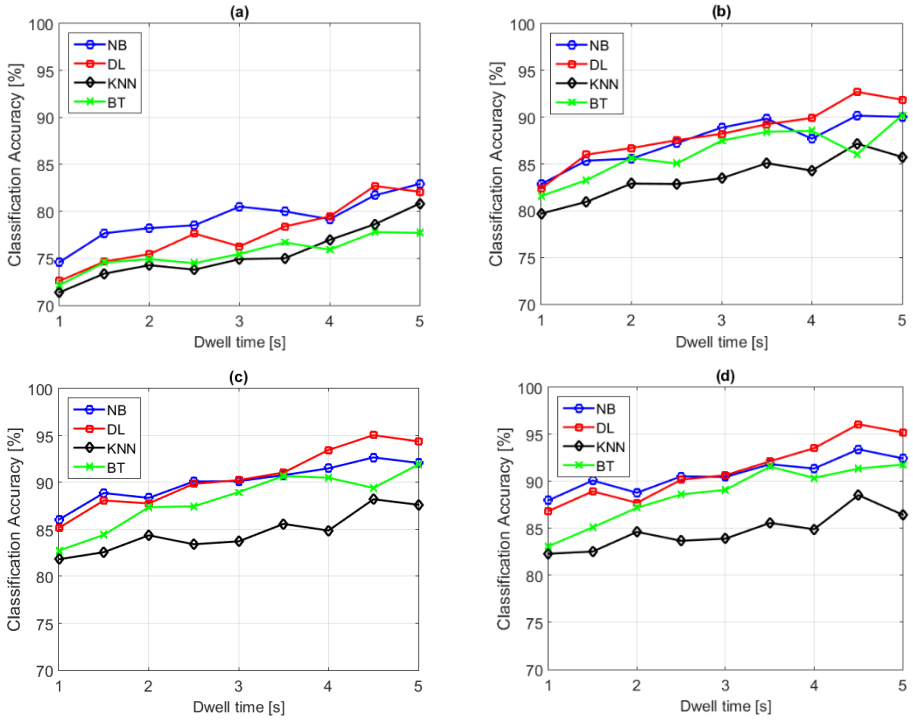

Figure 12 Classification accuracy vs dwell time using different classifiers and best combinations of features per radar node: (a) 1 feature, (b) 2 features, (c) 3 features, and (d) 4 features

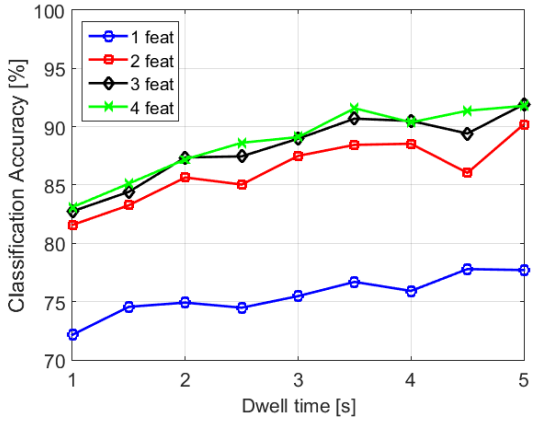

Figure 13 Classification accuracy vs dwell time and number of features at each radar node using BT classifier

Table 4 shows the classification accuracy obtained by the different classifiers when a single feature is selected through the wrapper method and used at each radar node. The three numbers in the table indicate the features used at node 1, node 2, and node 3 respectively. Fig. 14 represents on histograms the features selected at each radar node across the considered dwell times and classifiers. It can be seen that feature 2 appears to be very significant for node 2, but less significant at node 1 and node 3, where features 3 and 4 appears to be more relevant. It is also interesting to observe that some features are not useful for this particular classification problem and are never selected at any node, for instance feature number 1 or number 5. It is interesting to observe that node 2 was actually operating in cross-polarisation (i.e. receiving

H-polarised data) and co-located with node 1 as shown in Fig. 2, so the diversity in feature selection between these two nodes appears to be related to the difference in polarisation, as the aspect angle to the target is the same.

Table 4 Classification accuracy vs dwell time using different classifiers and the best combinations of single feature per radar node.

Classification accuracy [%]		1 s	1.5 s	2 s	2.5 s	3 s	3.5 s	4s	4.5 s	5 s
NB	Accuracy	74.6	77.7	78.2	78.5	80.5	80.0	79.2	81.7	82.9
	Features	4-2-10	4-2-10	3-2-4	3-2-4	3-2-4	3-2-4	12-2-3	3-2-4	3-2-4
DL	Accuracy	72.6	74.7	75.5	77.6	76.3	78.4	79.5	82.7	82.1
	Features	10-2-4	4-2-3	4-2-3	3-2-4	3-2-4	3-2-4	4-2-3	4-2-10	3-2-4
KNN	Accuracy	71.4	73.4	74.3	73.8	74.9	75.0	77.0	78.6	80.8
	Features	4-2-2	4-2-3	2-2-4	4-2-3	4-2-3	4-2-3	4-2-3	4-2-10-	6-9-2
BT	Accuracy	72.2	74.6	74.9	74.5	75.5	76.7	75.9	77.8	77.7
	Features	3-2-4	4-2-3-	4-2-3	3-2-4-	2-2-3	4-2-3	4-2-3	6-2-12	2-10-6

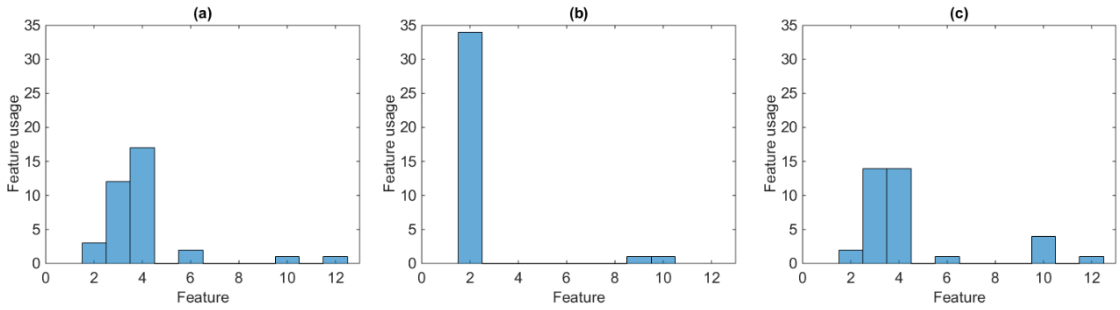

Figure 14 Histograms of features used at each radar node, with four classifiers and different dwell times: (a) node 1, (b) node 2, and (c) node 3

3.4 Additional analysis

This section presents additional analysis on the data related to the two classification scenarios described in this paper. The first test aims to investigate the effect of using a larger amount of the available data to train the chosen classifier, namely 70% of the available feature samples for training and the remaining for testing. A single feature for each node was identified using the wrapper approach and used for these examples. Fig. 15 shows the results for the BT classifier. Fig. 15a refers to the scenario for armed/unarmed classification and presents the accuracy as a function of dwell time for each aspect angle, similarly to what shown in Fig. 5d for 25% training. The trend of increasing accuracy with longer dwell time can be seen for all aspect angles, and in general the achieved accuracy is significantly higher using more data to train the classifier, even at the most unfavourable aspect angles, angle 4 and 5. The fact that the accuracy reaches almost 100%

for dwell time equal to or longer than 3s may be caused by a limited number of feature samples for testing, as only one feature sample per measurement can be obtained with these values of dwell time (i.e. each recording was only 5 s long). Fig. 15b compares the classification accuracy as a function of dwell time for the personnel recognition scenario. The BT classifier was used, trained with 25%, 50%, and 70% of the available feature samples. The trend of increasing accuracy with longer dwell time can be seen, as well as a significant improvement when the classifier was trained with more samples, which is up to between 5% and 10% when comparing 25% with 70% training.

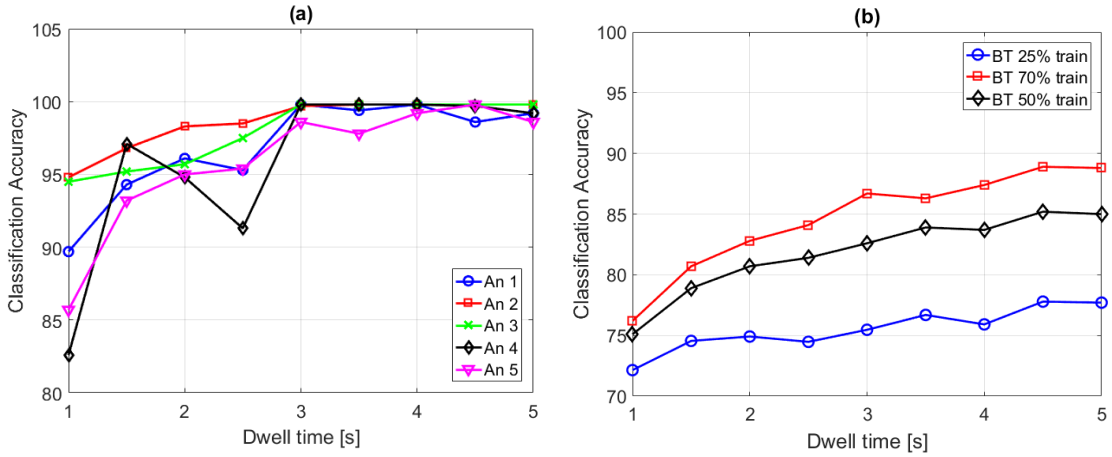

Figure 15 Classification accuracy vs dwell time using BT classifier trained with 70% of available data: (a) unarmed/armed classification scenario, and (b) personnel recognition scenario

The second test aims to present the performance of a more sophisticated type of classifier, the Support Vector Machine (SVM), and compare the classification accuracy with the simpler classifiers considered in the previous analysis [34, 39]. Both versions of SVM with linear kernel and with Radial Basis Functions (RBF) have been tested for the scenario of unarmed/armed classification, and the results are presented in Fig. 16a and Fig. 16b, respectively for linear and RBF kernels. The SVM classifiers were trained with 70% of the available data, so the results can be compared with those generated using the BT classifier in Fig. 15a. The trends in accuracy values as a function of dwell time appear similar for all classifiers, with lower values for shorter dwell times which then reach a plateau around 99% accuracy for dwell times longer than 3 s. Unfavourable aspect angles such as angle 4 and 5 present lower values of accuracy for shorter dwell times, but the accuracy appears to be consistently above 90% for dwell time above 2s. It is interesting to observe that the SVM classifier with RBF kernel outperforms the SVM classifier with linear kernel across

the considered aspect angles. Fig. 17 summarises the performance of both SVM classifiers and the BT classifier for aspect angle 1. The better performance of the RBF kernel version over the linear version can be seen, as well as the very similar performance of the SVM classifier with RBF kernel compared with the BT classifier, at least for the specific classification problem considered in this paper.

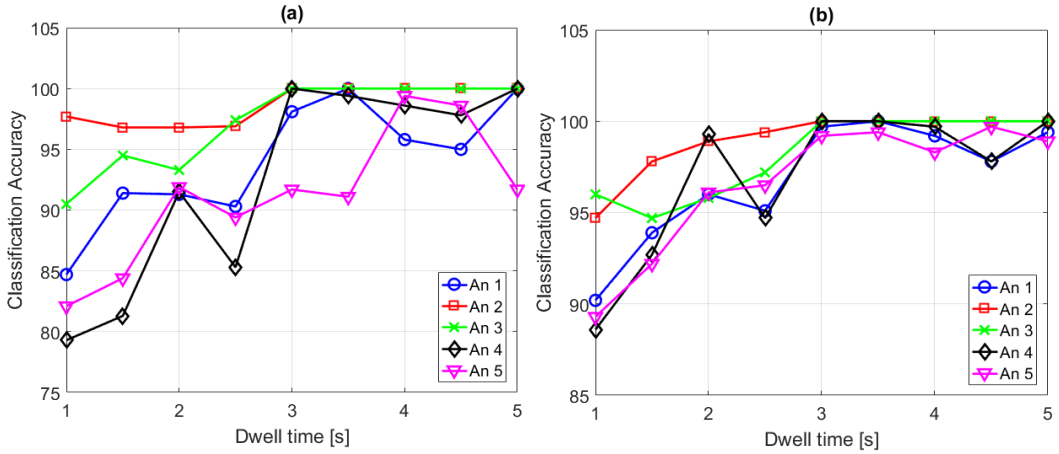

Figure 16 Classification accuracy vs dwell time using SVM classifier trained with 70% of available data: (a) linear kernel, and (b) RBF kernel

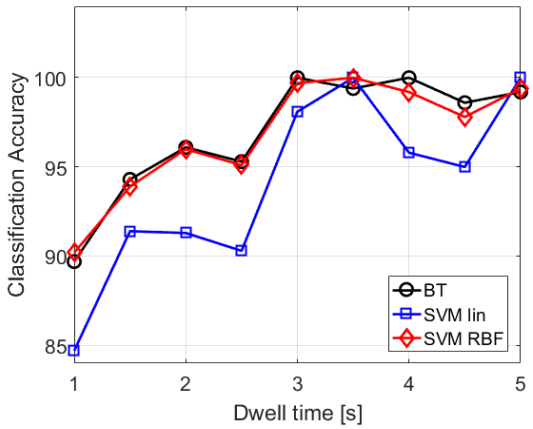

Figure 17 Classification accuracy vs dwell time for aspect angle 1 and classifiers trained with 70% of available data

A further test consists of assessing the classifier performance for armed/unarmed classification by using data from different subjects for the training and the testing steps. This test was performed for the DL, BT, and SVM with RBF kernel classifiers, trained with data from the two subjects who took part to the experiment in February 2016 as described in section 3.1, but tested with data previously collected in July 2015 where one of the subject was a different person. The deployment geometry of the radar nodes was the same as shown in Fig. 1a, but only a limited amount of data was collected with the subjects moving along aspect angle 1, i.e. walking straight towards the baseline. Three 5 s recordings were collected for the case

where both subjects were unarmed, and three for the case where only one of the two subjects was armed, hence a total number of 18 recordings considering all the three radar nodes. Fig. 18 shows the classification results for the three considered classifiers. Compared with the situation where the classifiers were trained and tested with data from identical targets, as in Fig. 5, the classification accuracy for BT and DL classifiers is reduced depending on the dwell time, up to the worst case scenario of a reduction of about 15%. The degradation is particularly evident for the BT classifier with short dwell times, but less significant for the DL classifier for which the accuracy is above 75%. A reduction in accuracy was expected because of the testing data from a new subject unknown to the classifier and because of the limited amount of data available. However, the SVM classifier appears to yield high level of accuracy, above 90%, comparable to the situation where the classifiers were trained and tested with data from the same subjects, as in Fig. 17. This seems to be an advantage of using a more complicated but more powerful classifier such as SVM in comparison with the simpler classifiers considered previously. The ability of the proposed features and classifiers to generalise well their performance even in the presence of data from new subjects is a very significant aspect for practical deployment. It is believed that the overall performance can be improved by collecting a larger database of data for training, involving more combinations of human subjects to capture the diversity in terms of body parameters such as height and weight and the different walking styles. This will be considered in future work to expand the results presented here.

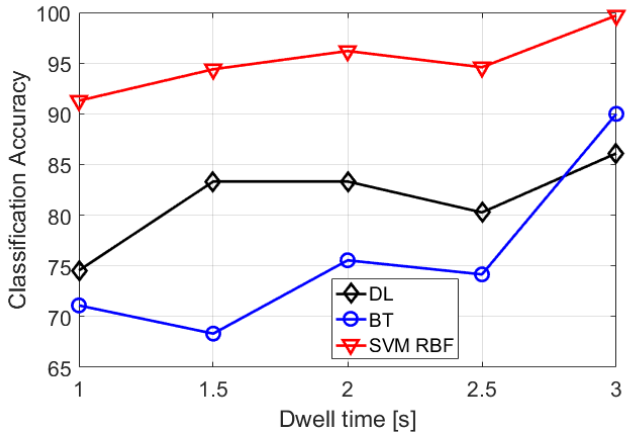

Figure 18 Classification accuracy for DL and BT classifiers tested and trained with data from different subjects

Finally, the computational complexity of the classifiers tested in this section is investigated in terms of processing time. The armed/unarmed classification scenario was considered, with 70% of the feature

samples used for training, 1 s dwell time, and data related to aspect angle 1, as from Fig. 1a. The best single feature at each node as identified by the wrapper approach was used. The classifiers were implemented on a standard desktop computer in MATLAB and tested in the same conditions. The results are summarised in table 5. The fastest classifiers appear to be the Binary Tree and the Nearest Neighbours, followed by the Diagonal-Linear and both SVM versions, and then by the Naïve Bayes. Overall, the differences in computational time do not appear to be very significant, but this may change if these algorithms were implemented on a different system with constrained computational resources.

Table 5 Computational efficiency for different classifiers

Processing Time [s] for each classifier	**NB**	**DL**	**KNN**	**BT**	**SVM linear**	**SVM RBF**
	0.783	0.715	0.643	0.639	0.724	0.712

4. Conclusions

This paper has investigated the performance gains possible through the exploitation of feature diversity at each node in a multistatic radar network. Two sets of data collected in different experiments for two different classification problems were specifically analysed; namely, classification of unarmed vs potentially armed personnel when two subjects are present together in the micro-Doppler signature, and personnel recognition of four different subjects based on their walking gait. Twelve different features based on the SVD and the centroid of the signatures have been considered together with four classifiers. These were chosen out of the many possible features proposed in the literature, as they can be easily and automatically extracted from the micro-Doppler signatures. Feature selection approaches based on brute-force wrapper and on ranking the features with a chosen metric (filter) have been compared. The results show that the best classification accuracy can be achieved by selecting different features at each radar node, and there is a significant influence of parameters such as dwell time and aspect angle on what features are most suitable. It is also shown that the conventional approach of having all the nodes selecting the same features leads to a decrease in performance (for instance the performance of a multistatic system for armed/unarmed personnel classification was shown to improve by as much as 15% in some cases by taking into account feature diversity at each node). This diversity in features providing the best classification accuracy was observed both in the first scenario for armed/unarmed classification and in the second scenario

for personnel recognition. In the former case all the three radar nodes operated with the same vertical polarisation, so it would appear that this behaviour is related to the spatial deployment of each node, which sees the target with a different aspect angle. In the latter case, two nodes (Node 1 and 2) were co-located but they received different polarisations, i.e. one vertical polarisation co-polarised with the transmitter, the other horizontal polarisation cross-polarised with the transmitter. In this case it would appear that the difference in features selection for processing the same data is related to the difference in co-polarised versus cross-polarised micro-Doppler signatures. It is felt that additional work is needed for further understanding of the effect of aspect angle and polarisation on the selection of optimal features, in order to see whether the trends highlighted in these data are confirmed.

Moreover, it is shown that while some features are never selected, others are consistently chosen. This choice varies depending upon classification problem and scenario, thus motivating the need for feature sets to be chosen dynamically. Multistatic nodes can potentially operate with different radar parameters, such as dwell time, polarisation, or even frequency band. By adapting node behaviour on not just receive but also transmit, it is anticipated that future technological development of cognitive radar systems will provide nodes able to change their feature extraction and selection scheme based on the environment conditions and target behaviour. For example, the classifier implemented within each radar node could have some form of base of knowledge (i.e. a sort of memory) with details on the most suitable set of features to extract and select, based on the information about the target aspect angle provided by the detection and tracking processes, either internally performed by the node itself or given to the node as external information. This base of knowledge can be generated during the training phase of the classifier, using both experimental data or data from suitable kinematic models of targets [40], and can be updated during the lifetime of the radar node while processing more and more target data progressively.

Future work aims at collecting additional data for the two classification problems analysed here, in order to verify the trends observed in the feature selection as a function of the various parameters, and to investigate in more details the effect of polarisation diversity. Additional features based on different processing of the signatures can be also considered, together with discarding those features that appeared to be less suitable from the analysis in this paper. Data from different subjects with different body

parameters and walking style will also be collected to investigate how the proposed classification approach can be generalised and become more robust when dealing with data from new subjects.

Acknowledgement

The authors are grateful to A. Amiri and J. Patel for their help in the field experiments. This work has been funded by the IET A F Harvey Prize 2013, awarded to Prof Hugh Griffiths. S. Gürbüz acknowledges the support of the Marie Curie International Reintegration Grant EU FP7 PIRG-GA-2010-268276 to support this work.

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, Jan. 2006.

[2] V. C. Chen, *The Micro-Doppler Effect in Radar*, Artech House, Boston/London, 2011.

[3] V. C. Chen, D. Tahmoush, W. J. Miceli, *Radar Micro-Doppler Signatures: Processing and Applications*, Institution of Engineering and Technology, 2014.

[4] L. Stankovic, S. Stankovic, I. Orovic, *Time Frequency Analysis of Micro-Doppler Signals Based on Compressive Sensing*, CRC Press, 2014.

[5] D. Tahmoush, 'Review of micro-Doppler signatures', *IET Radar, Sonar & Navigation*, vol. 9 (9), pp. 1140-1146, December 2015.

[6] K. Youngwook, H. Sungjae, K. Jihoon, 'Human detection using Doppler radar based on physical characteristics of targets', *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 289-293, 2015.

[7] F. Gini, M. Rangaswamy, *Knowledge-based Radar Detection, Tracking, and Classification*, John Wiley & Sons, 2008.

[8] K. Youngwook, L. Hao, ‘Human activity classification based on micro-Doppler signatures using a Support Vector Machine’, *IEEE Transactions on Geoscience and Remote Sensing,* vol. 47, pp. 1328-1337, 2009.

[9] R. M. Narayanan, M. Zenaldin, 'Radar micro-Doppler signatures of various human activities', *IET Radar, Sonar & Navigation*, vol. 9, (9), p. 1205-1215, December 2015.

[10] Y. Kim T. Moon, ‘Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8-12, January 2016.

[11] S. Z. Gürbüz, B. Erol, B. Çağlıyan, B. Tekeli, 'Operational assessment and adaptive selection of micro-Doppler features', *IET Radar, Sonar & Navigation*, vol. 9 (9), p. 1196-1204, December 2015.

[12] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. Soraghan, S. Marshall, ‘Robust PCA for micro-doppler classification using SVM on embedded systems’, *IEEE Transactions on Aerospace and Electronic Systems*, vol.50, (3). pp. 2304-2310, July 2014.

[13] B. Tekeli, S. Z. Gurbuz. M. Yuksel, ‘Information-theoretic feature selection for human micro-Doppler signature classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2749-2762, May 2016.

[14] R. Ricci, A. Balleri, ‘Recognition of humans based on radar micro-Doppler shape spectrum features’, *IET Radar, Sonar & Navigation*, vol. 9 (9), pp. 1216-1223, December 2015.

[15] F. Fioranelli, M. Ritchie, 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, May 2016.

[16] S. Björklund, H. Petersson and G. Hendeby, ‘On distinguishing between human individuals in micro-Doppler signatures’, *14th International Radar Symposium,* Dresden, June 2013, pp. 865-870.

[17] M. G. Amin, F. Ahmad, Y. D. Zhang and B. Boashash, 'Human gait recognition with cane assistive device using quadratic time–frequency distributions', *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1224-1230, 12 2015.

[18] D. P. Fairchild, 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, June 2014.

[19] A. Brewster, A. Balleri, 'Extraction and analysis of micro-Doppler signatures by the Empirical Mode Decomposition', *2015 IEEE Int. Radar Conference*, Arlington, VA, 2015, pp. 0947-0951.

[20] C. Karabacak, S. Z. Gürbüz, M. B. Guldogan, A. C. Gürbüz, 'Multi-aspect angle classification of human radar signatures', *Proc. SPIE 8734, Active and Passive Signatures IV*, 873408, May 23, 2013.

[21] B. Tekeli, S. Z. Gurbuz, M. Yuksel, A. C. Gurbuz, M. B. Guldogan, 'Classification of human micro-Doppler in a radar network', *2013 IEEE Radar Conference*, Ottawa, Canada, pp. 1-6, May 2013.

[22] G. E. Smith, K. Woodbridge, C. J. Baker, and H. Griffiths, 'Multistatic micro-Doppler radar signatures of personnel targets', *IET Signal Processing*, vol. 4, pp. 224-233, 2010.

[23] D. P. Fairchild, R. M. Narayanan, 'Multistatic micro-Doppler radar for determining target orientation and activity classification', *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 1, pp. 512-521, February 2016.

[24] F. Fioranelli, M. Ritchie, H. Griffiths, 'Multistatic human micro-Doppler classification of armed/unarmed personnel', *IET Radar, Sonar & Navigation*, vol. 9 (7), pp. 857-865, August 2015.

[25] F. Fioranelli, M. Ritchie, H. Griffiths, 'Aspect angle dependence and multistatic data fusion for micro-Doppler classification of armed/unarmed personnel', *IET Radar, Sonar & Navigation*, vol. 9 (9), pp. 1231-1239, December 2015.

[26] F. Fioranelli, M. Ritchie, H. Griffiths, 'Classification of unarmed/armed personnel using the NetRAD multistatic radar for micro-Doppler and Singular Value Decomposition features*', IEEE Geoscience and Remote Sensing Letters*, vol.12, no.9, pp.1933,1937, Sept. 2015.

[27] F. Fioranelli, M. Ritchie, H. Griffiths, ''Centroid features for classification of armed/unarmed multiple personnel using multistatic human micro-Doppler', *IET Radar, Sonar & Navigation*, accepted for publication in April 2016.

[28] S. Björklund, H. Petersson, G. Hendeby, 'Features for micro-Doppler based activity classification', *IET Radar, Sonar & Navigation*, vol. 9 (9), p. 1181-1187, December 2015.12 2015.

[29] S. Z. Gürbüz, B. Tekeli, M. Yüksel, C. Karabacak, A. C. Gürbüz and M. B. Guldogan, 'Importance ranking of features for human micro-Doppler classification with a radar network', *16th International Conference on Information Fusion*, July 2013, Istanbul, pp. 610-616.

[30] T. E. Derham, S. Doughty, K. Woodbridge, C. J. Baker, 'Design and evaluation of a low-cost multistatic netted radar system', *IET Radar, Sonar & Navigation*, vol. 1, pp. 362-368, 2007.

[31] F. Fioranelli, M. Ritchie, H. Griffiths, H. Borrion, 'Classification of loaded/unloaded micro-drones using multistatic radar', *Electronics Letters*, vol. 51, no. 22, pp. 1813-1815, 10 22 2015.

[32] A. Balleri, A. Al-Armaghany, H. Griffiths, K. Tong, T. Matsuura, T. Karasudani, *et al.*, 'Measurements and analysis of the radar signature of a new wind turbine design at X-band', *IET Radar, Sonar & Navigation*, vol. 7, pp. 170-177, 2013.

[33] J. J. M. De Wit, R. Harmanny, P. Molchanov, 'Radar micro-Doppler feature extraction using the Singular Value Decomposition', *2014 International Radar Conference*, Lille, France, 2014.

[34] T. Hastie, R. Tibshirani, J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*, Springer, 2009.

[35] C. D. Manning, P. Raghavan, M. Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.

[36] B. Frénay, G. Doquire, M. Verleysen, 'Theoretical and empirical study on the potential inadequacy of mutual information for feature selection in classification', *Neurocomputing*, vol. 112, July 2013, pp. 64-78.

[37] S. Theodoridis, K. Koutroumbas, *Pattern Recognition – Fourth edition*, Academic Press, 2009.

[38] C.E. Shannon, 'A mathematical theory of communication', *Bell Syst. Tech. Journal*, 27, 379-423 623-656, 1948.

[39] N. Cristianini, J. Shawe-Taylor, *An introduction to Support Vector Machines and Other Kernel-based Learning Methods*, Cambridge University Press, 2000.

[40] F. Gini, M. Rangaswamy, *Knowledge Based Radar Detection, Tracking and Classificaiton*, John Wiley Press, 2008.

Figures and Tables

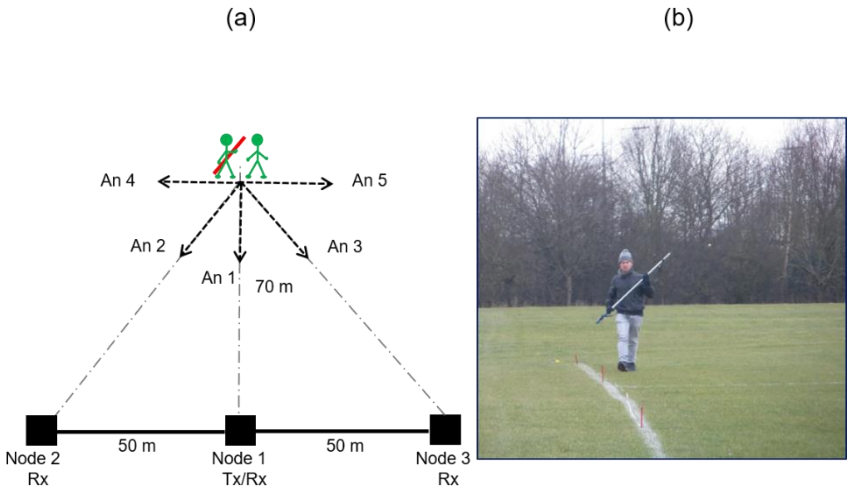

Figure 19 Measurement setup for unarmed vs armed classification experiment (a), and example of person carrying the pole representing the rifle (b)

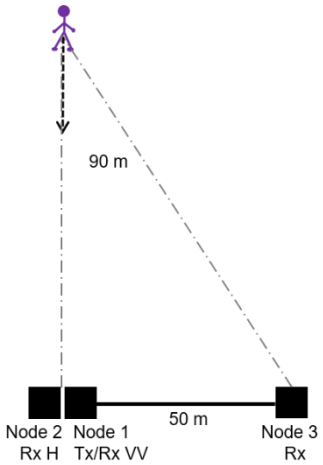

Figure 20 Measurement setup for personnel recognition experiment

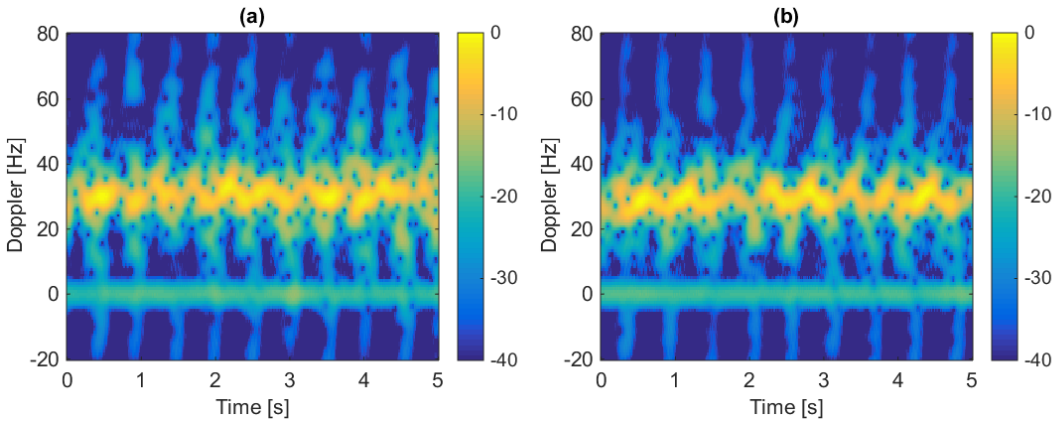

Figure 21 Spectrograms recorded at the monostatic node for two people walking together: both unarmed (a) and one armed and one unarmed (b)

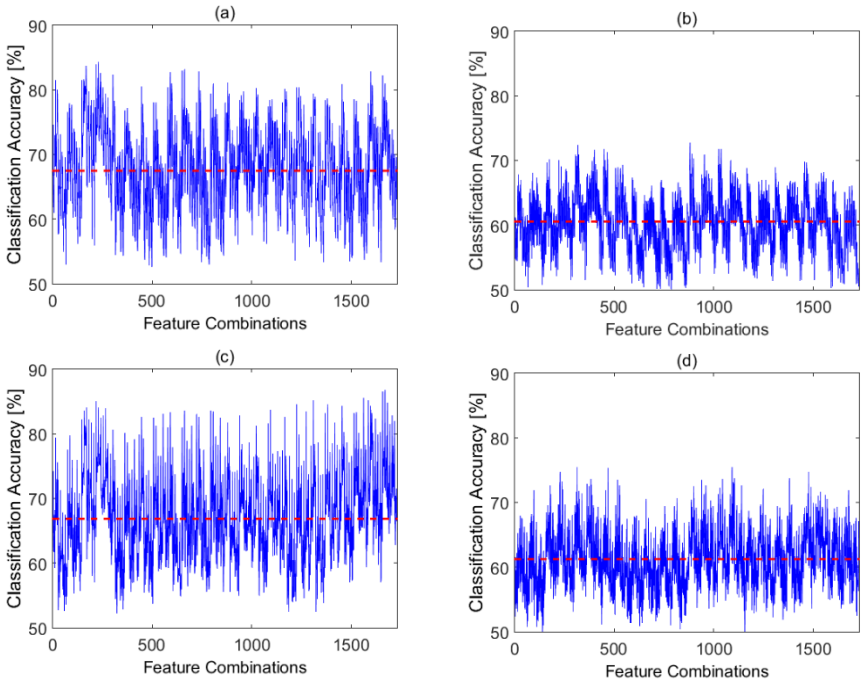

Figure 22 Classification accuracy vs different combinations of single features used at each radar node: (a) angle 1 and dwell time 1 s, (b) angle 4 and dwell time 1 s, (c) angle 1 and dwell time 2.5 s, and (d) angle 4 and dwell time 2.5 s

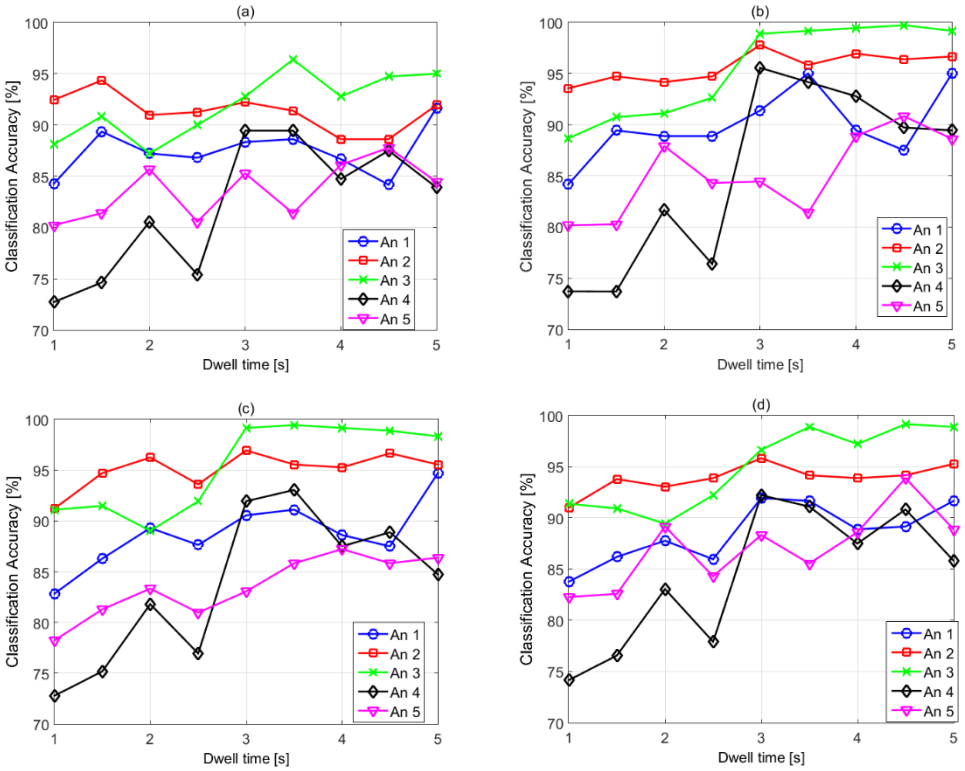

Figure 23 Classification accuracy vs dwell time using the best single feature at each radar node: (a) NB classifier, (b) DL classifier, (c) KNN classifier, and (d) BT classifier

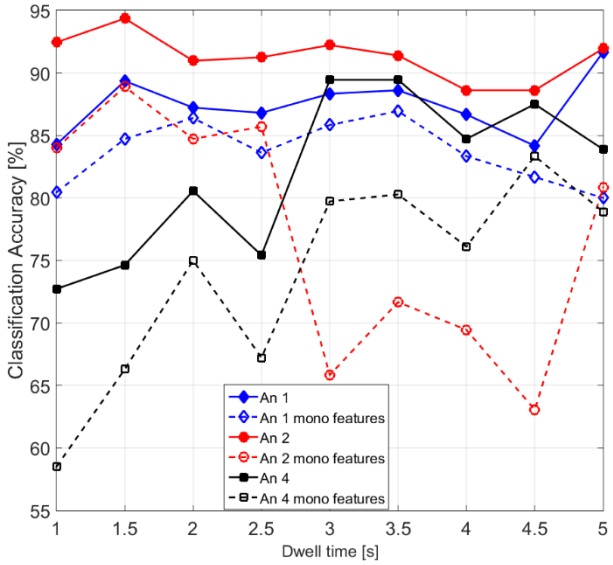

Figure 24 Classification accuracy comparison using best feature combination selected by wrapper approach vs using best combination for monostatic node at all radar nodes (indicated as 'mono features'). NB classifier with single feature per radar node was used.

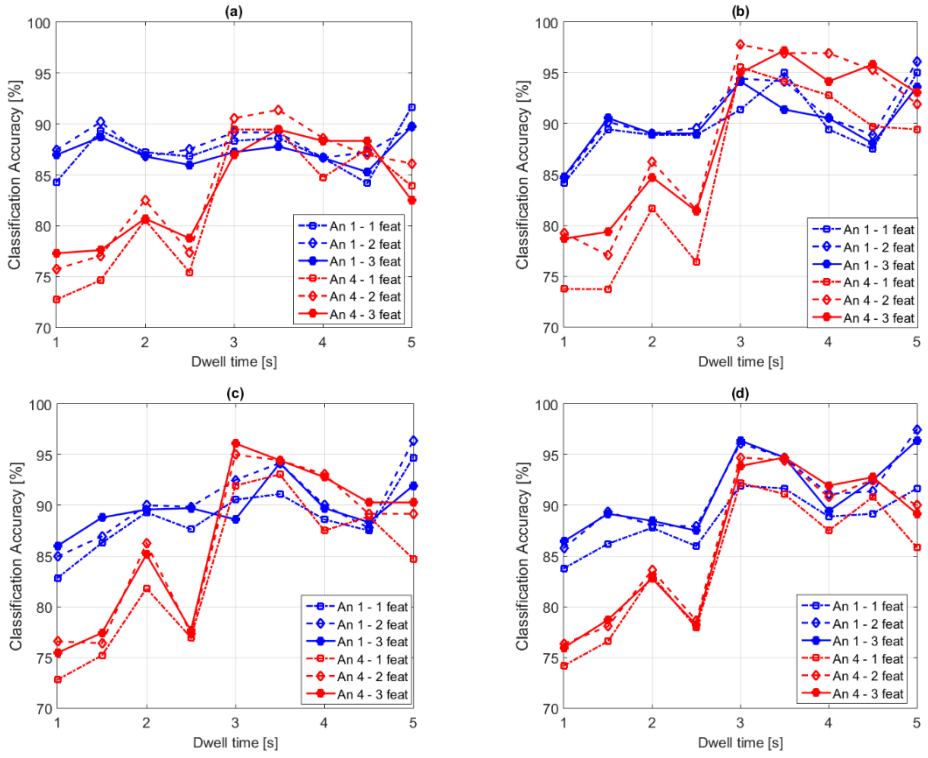

Figure 25 Classification accuracy vs dwell time using the best combination of 1 feature, 2 features, and 3 features per radar node: (a) NB classifier, (b) DL classifier, (c) KNN classifier, and (d) BT classifier

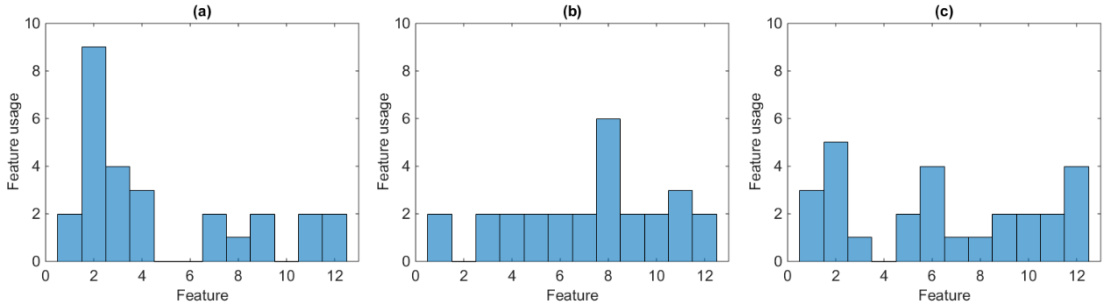

Figure-26 Histograms of features used at each radar node for aspect angle 1, BT classifier, and different dwell times: (a) node 1, (b) node 2, and (c) node 3

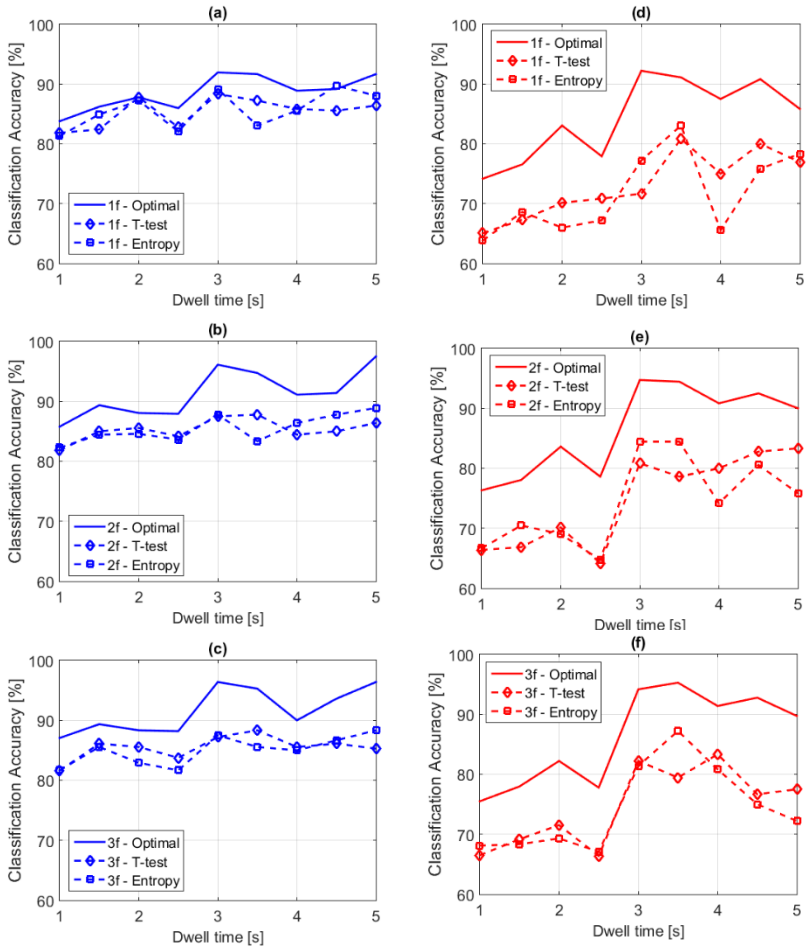

Figure 27 Classification accuracy vs dwell time using BT classifier and different combinations of features per node, selected by wrapper and ranking approaches: (a) 1 feature angle 1, (b) 2 features angle 1, (c) 3 features angle 1, (d) 1 feature angle 4, (e) 2 features angle 4, and (f) 3 features angle 4

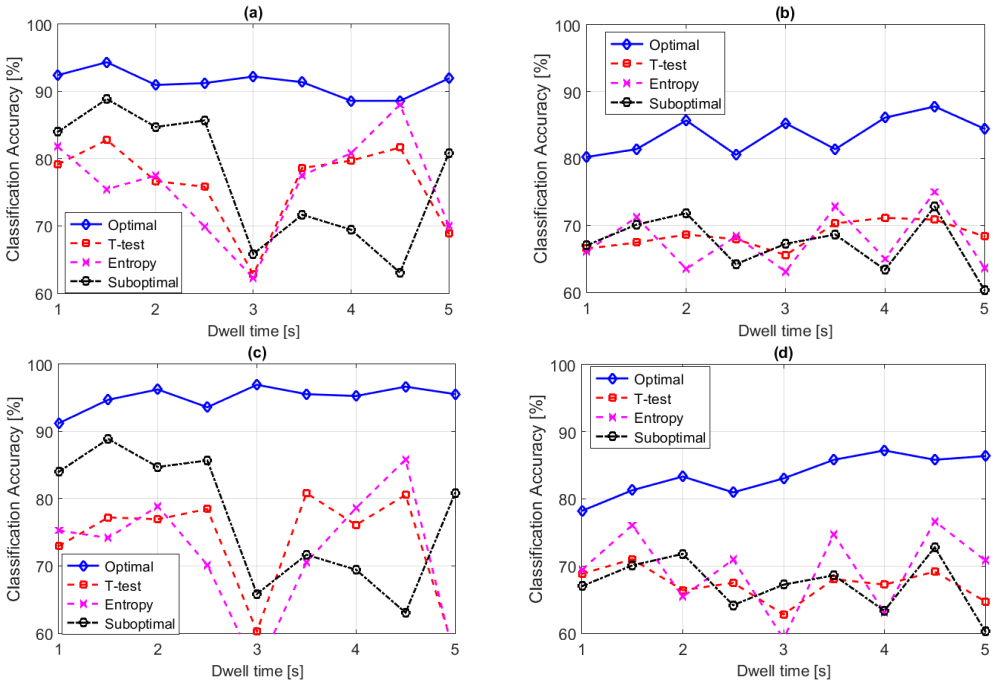

Figure 28 Classification accuracy vs dwell time using a single feature at each radar node selected by wrapper and ranking approaches: (a) angle 2 NB classifier, (b) angle 5 NB classifier, (c) angle 2 KNN classifier, and (d) angle 5 KNN classifier

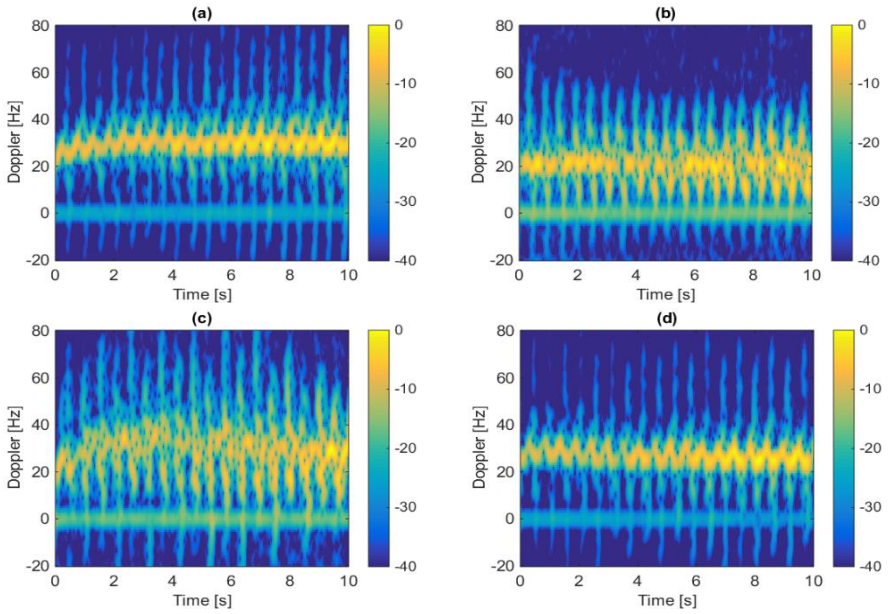

Figure 29 Spectrograms recorded at the monostatic node for different people walking: person 1 (a), person 2 (b), person 3 (c), and person 4 (d)

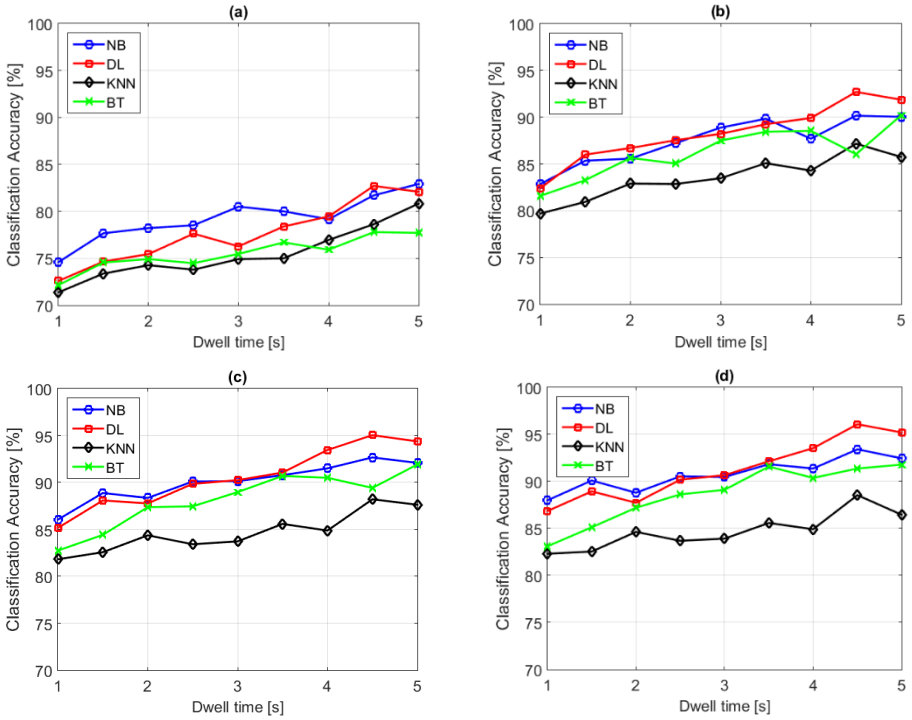

Figure 30 Classification accuracy vs dwell time using different classifiers and best combinations of features per radar node: (a) 1 feature, (b) 2 features, (c) 3 features, and (d) 4 features

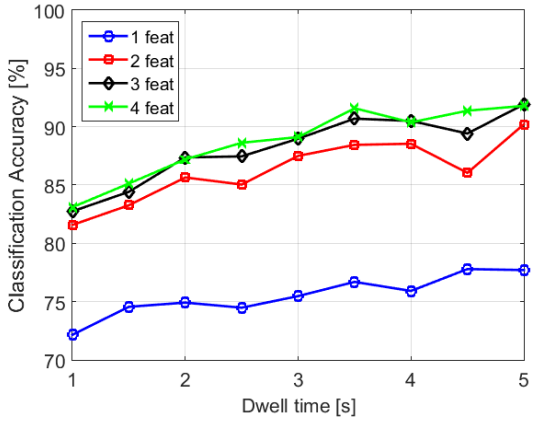

Figure 31 Classification accuracy vs dwell time and number of features at each radar node using BT classifier

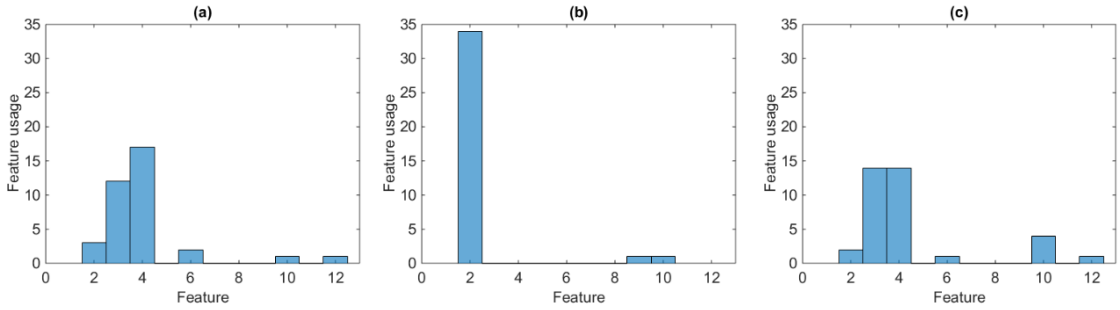

Figure 32 Histograms of features used at each radar node, with four classifiers and different dwell times: (a) node 1, (b) node 2, and (c) node 3

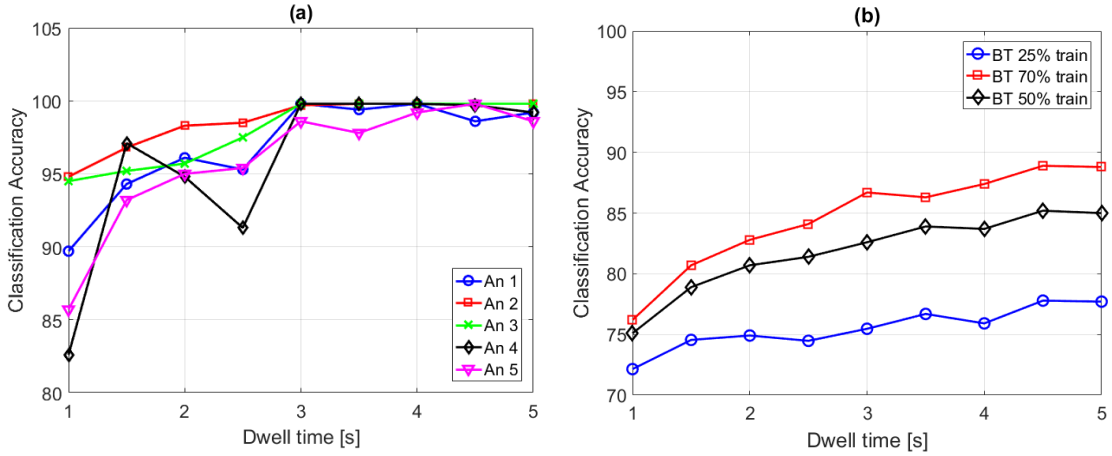

Figure 33 Classification accuracy vs dwell time using BT classifier trained with 70% of available data: (a) unarmed/armed classification scenario, and (b) personnel recognition scenario

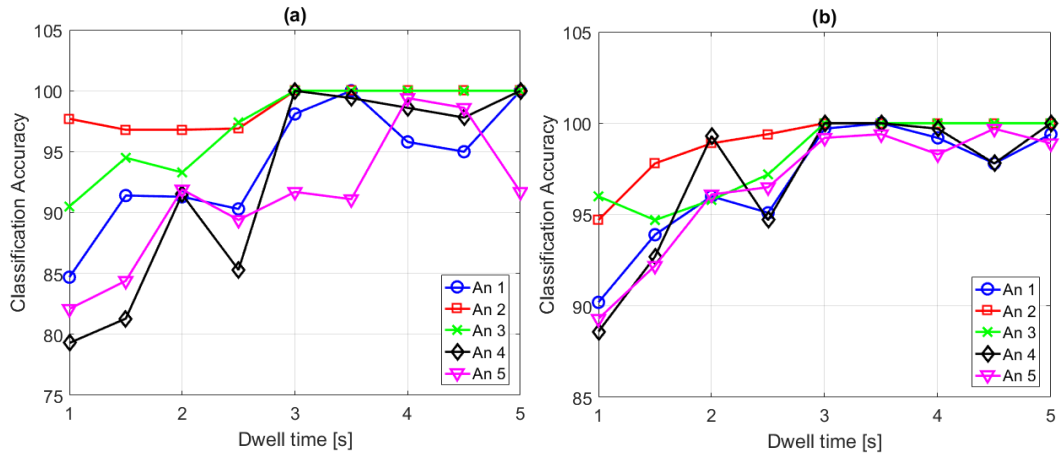

Figure 34 Classification accuracy vs dwell time using SVM classifier trained with 70% of available data: (a) linear kernel, and (b) RBF kernel

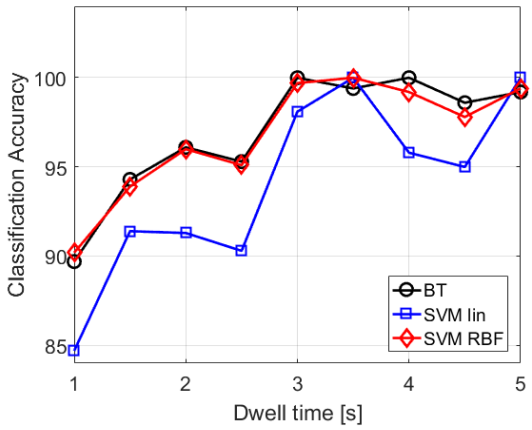

Figure 35 Classification accuracy vs dwell time for aspect angle 1 and classifiers trained with 70% of available data

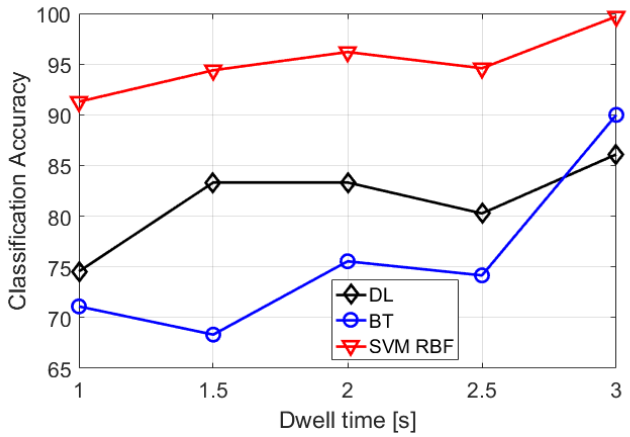

Figure 36 Classification accuracy for DL and BT classifiers tested and trained with data from different subjects

Table 6 Classification accuracy vs dwell time and aspect angle using BT classifier. The single feature used at each radar node is also indicated.

Classification accuracy [%]		**1 s**	**1.5 s**	**2 s**	**2.5 s**	**3 s**	**3.5 s**	**4s**	**4.5 s**	**5 s**
An 1	**Accuracy**	83.8	86.2	87.8	86.0	91.9	91.7	88.9	89.2	91.7
	Features	2-8-5	2-11-5	2-2-11	2-2-10	2-9-12	2-10-2	10-9-2	2-1-2	2-11-10
An 2	**Accuracy**	91.0	93.8	93.1	93.9	95.8	94.2	93.9	94.2	95.3
	Features	1-8-2	1-2-2	1-2-2	1-12-2	9-1-2	1-2-2	9-1-2	9-1-2	1-4-2
An 3	**Accuracy**	91.4	90.9	89.4	92.2	96.7	98.9	97.2	99.2	98.9
	Features	1-3-11	1-11-11	1-2-2	4-11-11	7-2-2	8-2-2	8-3-2	1-1-8	1-11-1
An 4	**Accuracy**	74.2	76.6	83.1	77.9	92.2	91.1	87.5	90.8	85.8
	Features	4-3-3	10-7-12	5-11-2	10-7-2	8-12-3	7-8-2	7-12-2	2-12-2	13-3-3
An 5	**Accuracy**	82.3	82.6	89.2	84.3	88.3	85.6	88.6	93.9	88.9
	Features	3-11-2	3-11-2	3-11-2	12-7-2	3-4-4	4-2-3	10-3-2	11-1-3	11-12-2

Table 7 Classification accuracy vs dwell time using BT classifier and the best combinations of three features per radar node.

Classification accuracy [%]		**1 s**	**1.5 s**	**2 s**	**2.5 s**	**3 s**	**3.5 s**	**4s**	**4.5 s**	**5 s**
An 1	**Accuracy**	87.1	89.4	88.3	88.2	96.4	95.3	90.0	93.6	96.4
	Features N1	2-11-1	2-9-3	2-7-4	2-4-12	2-7-9	2-3-12	10-1-4	2-3-8	2-3-11
	Features N2	8-1-7	11-1-5	2-8-6	2-11-7	9-10-3	10-4-5	9-12-3	1-4-6	11-12-8
	Features N3	5-2-11	5-11-6	11-2-3	10-11-6	12-5-7	2-10-9	2-10-1	2-6-8	10-6-5
An 4	**Accuracy**	75.5	78.0	82.2	77.8	94.2	95.3	91.4	92.8	89.7
	Features N1	4-6-5	10-3-6	5-10-2	10-3-2	8-2-7	7-12-6	7-8-9	2-9-5	12-2-8
	Features N2	3-2-5	7-5-2	11-4-5	7-8-12	12-8--11	8-5-12	12-8-5	12-5-1	3-6-7
	Features N3	3-10-11	12-3-9	2-3-1	2-1-4	3-11-2	2-3-6	2-11-7	2-8-11	3-8-2

Table 8 Classification accuracy vs dwell time using BT classifier and single feature selected at each node using wrapper and ranking approaches. Results related to aspect angle 4.

Classification accuracy [%]		1 s	1.5 s	2 s	2.5 s	3 s	3.5 s	4s	4.5 s	5 s
Optimal	Accuracy	74.2	76.6	83.1	77.9	92.2	91.1	87.5	90.8	85.8
	N1	4	10	5	10	8	7	7	2	12
	N2	3	7	11	7	12	8	12	12	3
	N3	3	12	2	2	3	2	2	2	3
T-test	Accuracy	65.1	67.3	70.1	70.8	71.7	80.8	75.0	80.0	76.9
	N1	3	2	2	12	2	12	12	2	12
	N2	10	3	11	3	3	11	3	3	3
	N3	3	3	3	3	3	3	2	3	2
Entropy	Accuracy	63.8	68.6	66.0	67.2	77.2	83.1	65.6	75.8	78.3
	N1	2	2	5	12	2	12	12	12	12
	N2	2	2	11	3	3	11	6	6	4
	N3	10	10	3	3	2	2	2	2	2

Table 9 Classification accuracy vs dwell time using different classifiers and the best combinations of single feature per radar node.

Classification accuracy [%]		1 s	1.5 s	2 s	2.5 s	3 s	3.5 s	4s	4.5 s	5 s
NB	Accuracy	74.6	77.7	78.2	78.5	80.5	80.0	79.2	81.7	82.9
	Features	4-2-10	4-2-10	3-2-4	3-2-4	3-2-4	3-2-4	12-2-3	3-2-4	3-2-4
DL	Accuracy	72.6	74.7	75.5	77.6	76.3	78.4	79.5	82.7	82.1
	Features	10-2-4	4-2-3	4-2-3	3-2-4	3-2-4	3-2-4	4-2-3	4-2-10	3-2-4
KNN	Accuracy	71.4	73.4	74.3	73.8	74.9	75.0	77.0	78.6	80.8
	Features	4-2-2	4-2-3	2-2-4	4-2-3	4-2-3	4-2-3	4-2-3	4-2-10-	6-9-2
BT	Accuracy	72.2	74.6	74.9	74.5	75.5	76.7	75.9	77.8	77.7
	Features	3-2-4	4-2-3-	4-2-3	3-2-4-	2-2-3	4-2-3	4-2-3	6-2-12	2-10-6

Table 10 Computational efficiency for different classifiers

Processing Time [s] for each classifier	NB	DL	KNN	BT	SVM linear	SVM RBF
	0.783	0.715	0.643	0.639	0.724	0.712

Francesco Fioranelli received the PhD degree from Durham University, UK, in 2014 and worked as Research Associate at the UCL Radar Group with Prof H. Griffiths between February 2014 and March 2016. He is currently a Lecturer at the School of Engineering at the University of Glasgow.

Matthew Ritchie received the D.Eng. degree from University College London (UCL), London, U.K., in association with Thales U.K., in 2013. He is currently a Research Associate with the UCL Radar Group, working with Prof. H. Griffiths.

Sevgi Z. Gürbüz received the Ph.D. in Electrical and Computer Engineering from the Georgia Institute of Technology, Atlanta, GA, in December 2009, the M.Eng. in Electrical Engineering and Computer Science in 2000, and the B.S. in Electrical Engineering with minor in Mechanical Engineering in 1998, both from the Massachusetts Institute of Technology, Cambridge, MA. She is currently a Research Scientist in the Department of Electrical and Computer Engineering at Utah State University, Logan, UT.

Hugh Griffiths received the M.A. degree in physics from Keble College, University of Oxford, Oxford, U.K., in 1978 and the Ph.D. and D.Sc. (Eng.) degrees from the University of London, London, U.K., in 1986 and 2000, respectively. He holds the THALES/Royal Academy of Engineering Chair of RF Sensors at University College London, London. He has published over 500 papers and technical articles in his areas of interest. He also serves on the Defence Scientific Advisory Council for the U.K. Ministry of Defence. He is a Fellow of The Institution of Engineering and Technology, and in 1997, he was elected to the Fellowship of the Royal Academy of Engineering.