Multi-domains based Human Activity Classification in Radar

*Zhenghui Li[1], Francesco Fioranelli[2], Shufan Yang[1], Lei Zhang[1], Olivier Romain[3], Qian He[4], Guolong Cui[4], Julien Le Kernec[1,3,4]**

[1]*James Watt School of Engineering, University of Glasgow, Glasgow, UK*
[2]*MS3-Microwave Sensing Signals and Systems, TU Delft, Delft, The Netherlands*
[3]*ETIS – Signal and Information Processing lab, University Cergy-Pontoise, Cergy, France*
[4]*School of Information and Communication, University of Electronic Science and Technology of China, Chengdu, China*
**Julien.lekernec@glasgow.ac.uk*

Keywords: CLASSIFICATION, HUMAN ACTIVITY RECOGNITION, HUMAN MICRO-DOPPLER, MACHINE LEARNING.

Abstract

In human activity recognition (HAR) based on radar, significant research exists on statistical features extracted from the spectrogram (μD), whereas the research which considers other domains is less developed. This paper is aimed to investigate three domains of radar data: μD, Cadence Velocity Diagram (CVD), and range-time (RT) information, evaluating which ones are best suited to classify specific activities. In addition, information fusion is applied to enhance classification accuracy and compare it with the results of single domain approach. Based on the previous results, a hierarchical structure is proposed to improve the performance of classification further. The preliminary results show that different domains have distinctive sensitivity to specific activities. RT information is sensitive to the moving target crossing range bins, while CVD is more sensitive to body movement. The μD is more balanced, which means it can observe both moving targets and body movements. Furthermore, improvement in accuracy is approximately 6-23 % using feature-level fusion. A hierarchical classification approach is also investigated, which has accuracy in the order of approximately 92 %.

I Introduction

Human Activity Recognition is required in various applications such as motion analysis and healthcare. Previously, a series of approaches have been proposed for activity classification [1, 2] based on wearable devices and optical devices, which has many limitations. Currently, radar is considered as a relevant technology in human activity recognition, with some unique advantages such as insensitivity to lighting and weather conditions, visual privacy protection, and safety compared with cameras and wearable devices.

The archetypal radar system transmits electromagnetic waves to the target, and then receives them when the waves are backscattered from the target. The range-time (RT) profile directly illustrates how the distance changes between target and radar over time. The apparent carrier frequency of radar will be shifted when the waves are reflected from a moving target, which is also called the Doppler effect [3]. Besides, some additional frequency modulations (micro-Doppler) generated by 'small movements' such as vibration and rotation, are added to the main Doppler. They appear around the torso main Doppler contribution due to micro-movements such as waving arms, swinging legs, and finger motion [3, 4]. The distinctive motions of different body parts constitute a specific micro-Doppler signature, and thus they can be employed in human activity recognition [5, 6].

Furthermore, the Cadence velocity diagram (CVD) is also a readily available approach that is obtained from the μD through a simple FFT along the time dimension for each Doppler bin [12]. Each domain has its own limitations. The μD and the CVD only present the variance of the velocity of different body parts, which completely neglects the range profile. The range profile is confusing to observe the 'small movements' of the human body. Thus, it is usually treated as a supplement tool in human activity recognition, with typically less developed and more limited analysis.

For the past few years, significant works have focused on the μD and its relative techniques [7]. In [10], the authors combined the range information and the micro-Doppler signature, with deep learning, improving the accuracy of the fall detection to approximately 98 %. In [11], the authors designed a new 3D model, range-velocity-time points, to describe micro-motions under multi-target conditions. Several classification techniques such as support vector machine and linear discriminant analysis were compared with the author's method in [8], and the conclusions that both support vector machine and Naïve Bayes algorithm are sufficient to distinguish micro-Doppler signatures of different activities. In [9], singular value decomposition (SVD) algorithm was applied to extract features from μDs. The author used a NetRAD system, which was developed at University College London, measuring human micro-Doppler signature. SVD was applied to the micro-Doppler signature, and features were extracted from the SVD matrix. The accuracy of the classification was approximately 99 %. In [13, 14], features from the CVD were used with data fusion techniques to be fused with features from other sensors or domains, and both get about 92 % accuracy.

Although the methods with other domains such as CVD are emerging, most of the researches still focuses on the μD [15].

The investigation of domains can provide more choices and techniques of cognitive selection, where not only the features but also data domains can be considered. The results can be implemented in other algorithms, such as hierarchical classification. Hierarchical classification is proposed to approach the problem by separating the activities into several subgroups, which can consider different domains and features for each subgroup classification. Dividing between those groups of activities would allow for both domain knowledge and information fusion approaches to be leveraged to further increase the overall accuracy, by pairing the radar data domain to the activities they are more suitable for.

In this work, three distinctive domain-based classifications, namely μD-based, CVD-based, and range-profile based, will be described and employed to the same human activity dataset. Comparing the classification results, the question of which type of activity will have a better performance in which domain will be answered. Furthermore, an information fusion and hierarchical approach will also be implemented to explore a better overall performance of classification.

The paper is organised as follows: in Section. II, the dataset and approaches are presented. Section. III deals with the feature extraction and classification problem and shows the results with discussion and improvements. Finally, Section IV will conclude this research.

II. Data Collection and Feature Extraction

The data analysed in this paper were collected using an off-the-shelf Frequency Modulated Continuous Wave (FMCW) radar sensor, which operates at a carrier frequency of 5.8 GHz, with 1 ms chirp duration and 400 MHz bandwidth. The output power of the radar is approximately +18 dBm. The radar is connected to two Yagi antennas, one for transmitting and the other for receiving, with a gain of about +17 dB.

A total of 1754 motion data files were recorded from 72 participants aged 21 to 98 years old. The overall dataset was composed of seven independent datasets; each dataset was collected in a different time and indoor environments. Six different types of daily activities (Table I) were recorded: walking, sitting down, standing up, pick up an object, drink water, and fall. Note that the dataset is not balanced, i.e., the number of activities per class was different, and some classes like fall were not performed by all participants due to research ethics and security problem (the elders could not be asked to performe fall).

Table I List of activities

No.	Activity Description
A1	Walking back and forth
A2	Sitting down on a chair
A3	Standing up from a chair
A4	Picking up an object from the ground
A5	Drinking water from a glass
A6	Fall

Extracting the salient features from the radar data guarantees the accuracy of the classification results. The extraction of features is inspired by [9, 16-18]. Fig. 1-3 indicates graphs of three domains of diverse activities performing by the same participant, a young male adult, whereas Fig. 4 is obtained from an older woman. For RT radar data, the features were extracted from the resulting images. Fig. 1 illustrates six examples of RT information. There are in total 21 features which are extracted from this domain, namely entropy, skewness, energy curve (mean, variance & root mean square), raw features from I and Q channel data [18] and the singular value decomposition (SVD) result of the RT information (mean and variance of the first three left vectors). The μDs (Doppler vs. time patterns) were calculated by applying Short Time Fourier Transform (STFT) to the raw RT radar data, where there are six examples demonstrating in Fig. 2. Besides, the radar dataset records activities from not only the adults but also the elders. Fig. 4 shows the micro-Doppler signature of different five activities of an older adult. (Note that there is no fall action due to security and ethics problem.) Compared with Fig. 2 (a), the differences caused by age can be observed clearly. The number of features extracted from the μD was 21, comprising entropy, skewness, bandwidth (mean & variance), centroid (mean & variance), energy curve and SVD results. The CVD is obtained from the μD through a simple FFT. The examples of CVD are shown in Fig. 3. The number of features of CVD is eight, which includes step repetition frequency, the energy of the main peak, peak velocity, intensity of the main peak, peak cadence frequency and band peak velocity.

The result of activity classification experiments will be demonstrated in Section III, with a quantitative comparison of results that are based on different domains. Besides, feature selection methods will be introduced to improve the performance of classification and reduce computational cost.

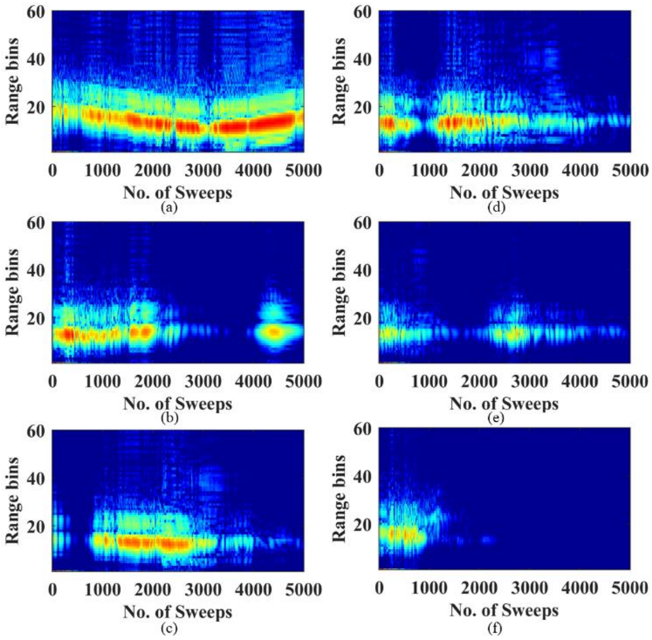

Fig. 1 The time range information of a young adult performing different activities. (a) walking, (b) sitting down, (c) standing up, (d) picking up an object, (e) drinking, (f) fall.

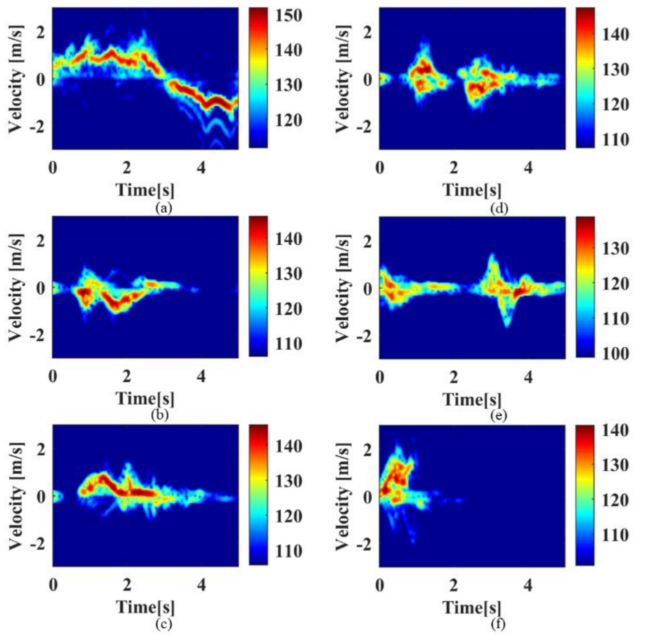

Fig. 2 The μD of different activities of a young adult. (a) walking, (b) sitting down, (c) standing up, (d) picking up an object, (e) drinking, (f) fall.

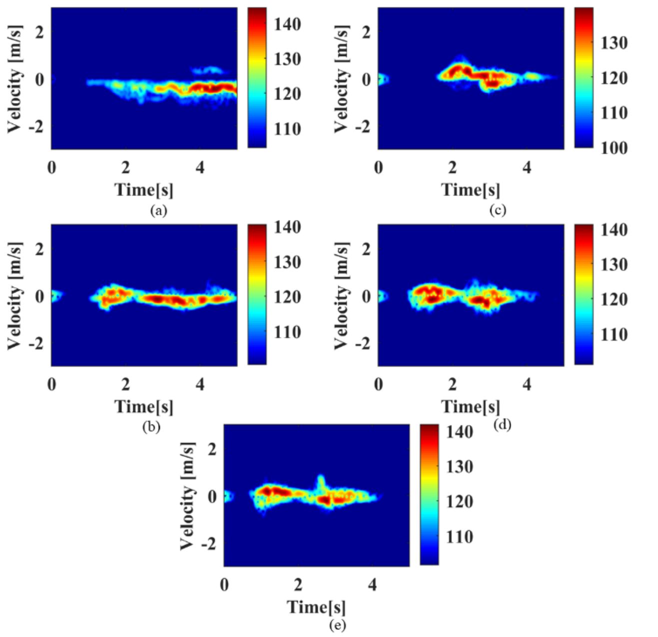

Fig. 4 The μD of activities of an older adult. (a) walking, (b) sitting down, (c) standing up, (d) picking up an object, (e) drinking

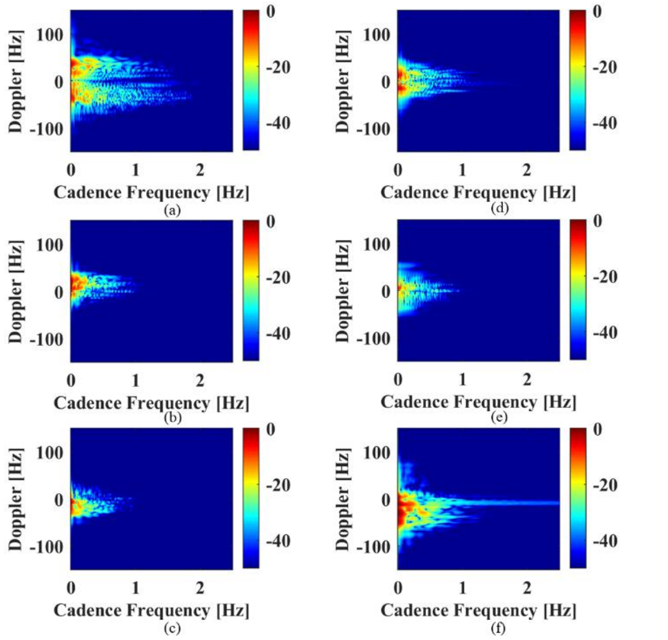

Fig 3 The CVD of different activities of a young adult. (a) walking, (b) sitting down, (c) standing up, (d) picking up an object, (e) drinking, (f) fall.

III Result

Based on the features that have been introduced in Section II, the classification results are presented first explaining the classification from one domain, classification from one domain by selecting features, and finally, from the fusion of information in different domains.

3.1 Classification

At this stage, an SVM classifier with a quadratic kernel and a weighted KNN classifier with K = 10 were employed to the activity classification. The validation of classification used k-fold cross-validation approach with k equal to ten.

Table II Comparison of classification accuracy of algorithms and domains

Accuracy (%)	μD	RT	CVD
SVM	80.3	64.2	**82.4**
KNN	75.2	61.3	**80.4**

Table II shows preliminary results of the classification without feature selection. Generally, the SVM algorithm is outperforming the KNN algorithm, and CVD obtains the best result overall. The result for μDs is worse than CVD by approximately 2 %. The classification performance of the RT domain is the worst, which is 18 % lower than CVD 18 % with SVM and 19 % with KNN.

3.2 Feature Selection

The purpose of feature selection is to investigate the optimisation of feature subset. Correctly, feature selection approaches are used for removing redundant or correlated features, referred to as confusing information, and thus it improves the classification accuracy and reduces computational load [16]. There are three main strategies of feature selection algorithms:

- Wrapper methods. It trains each subset and considers the different combinations of feature space with a specific classifier, using the error rate to find the result with the highest accuracy. For a particular model, the wrapper method usually has the best performance. However, it is computationally intensive due to that each subset needs to be trained.

- Filter methods. It evaluates the intrinsic relevance between features based on the metric of class separability [16], scoring the feature subset. High score features are selected while the low scores are discarded. Filter method is independent of the type of the classifier, and thus it is more general.
- Embedded methods. In this category, the feature selection techniques are integrated with the classification algorithm.

In this work, two filter-based methods, the Chi-Square [19] and Fisher score (F-score) [16], are investigated and implemented to select the optimal feature subsets from the feature pool. Chi-Square test is used in statistics for testing the independence of events. When it is used for feature selection, Chi-Square calculation indicates the dependence between the target and features, where higher the Chi-Square value, more informative the feature. Chi-Square is widely used due to its ease of computation and robustness with respect to the independence of data. Fisher score algorithm ranks the features using the distance, where the same class features have minimal distance, whereas the distance between different class features being maximal. Fisher score is also a computationally simple algorithm, with fast processing speed and generally good performance [16].

Table III Comparison of feature selection methods

Methods	#Features	Computing time (s)	Accuracy (%)
µD (SVM and FS)	13	41.8	81.9
µD (KNN and FS)	13	7.3	80.7
CVD (SVM and FS)	5	28.4	82.8
CVD (KNN and FS)	6	**1.7**	82.5
RT (SVM and FS)	18	126.4	65.1
RT (KNN and FS)	13	6.1	63.6
µD (SVM and CS)	15	56.5	**84.3**
µD (KNN and CS)	13	7.5	80.1
CVD (SVM and CS)	5	20.6	84.2
CVD (KNN and CS)	5	2.0	82.8
RT (SVM and CS)	17	97.2	67.6
RT (KNN and CS)	12	6.4	64.2

* Chi-Square: CS, Fisher Score: FS

Table III illustrates the results of the Fisher score (FS) and Chi-Square (CS). Fisher score method provides a limited improvement in the SVM algorithm, which is approximately 1 %, whereas it enhances the performance of KNN algorithm by approximately 3 %. The accuracy of the classification results is boosted when the CS is implemented. For the SVM result, it generally improves the accuracy by approximately 4 %, while this enhancement on the classification performance for KNN is about 5 % for the µD domain. Feature selection decreases the number of features to improve the computing time. FS shows that the optimisation occurs at 62-85 % feature available. Meanwhile, the optimal results obtained by CS happened around 57-80 % available features. Hence, the CS algorithm outperforms the FS approaches with higher accuracy of classification. Fig. 5 illustrates how the classification accuracy changes along with the increase of the number of features.

The purpose of this paper is to discuss and investigate the relationship between the domains and specific activities. Table IV summarises the results. Picking up an object is the most easily misclassified activity, where the highest accuracy is 77.3 %. In the µD domain, most misclassifications are from A4 (picking up an object), which average accuracy is 63.6 %. In the CVD domain, the most confusing pair is picking up an object and drinking water. The walking activity also has errors, while it performs well in the other two domains, with 100 % accuracy. However, the CVD domain has the best overall performance among the three domains, especially for its sensitivity to the sitting and standing, with high accuracies of 92.7 % and 89.0 %, respectively. In the RT domain, walking and fall can be well detected, with approximately 100 % and 87.1 % accuracy, whereas the rest of the activities are problematic.

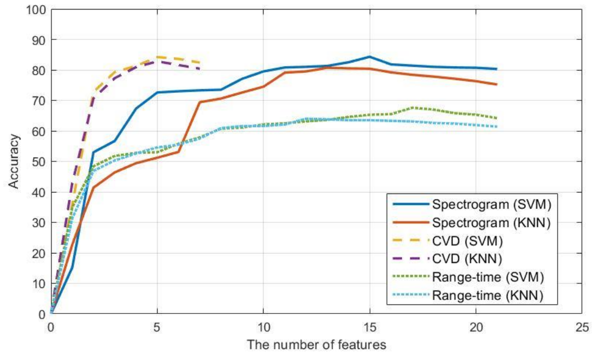

Fig. 5 Feature selection results using CS for three domains with SVM.

3.3 Data Fusion

To further investigate the overall performance of classification, data fusion approaches will be used, and the fusion results will be compared. Data fusion is the process of integrating multiple data sources to produce information which can overcome limitation caused by any single data source [20]. This can be achieved at many levels of abstraction, such as characteristics and symbols, and it is typically done at either signal, feature, or decision level [21]. In this paper, the feature level fusion is used on three different data domains. Feature level fusion is aimed to generate a single, larger feature vector samples from different features. Feature selection methods can be used for removing redundant or incorrect features in the new feature space. For feature level fusion, the Chi-Square algorithm is used as feature selection method before feeding features to the classifiers due to its higher accuracy, and the SVM classifier only, as it was found to be the better performing classifier compared with KNN.

The results are shown in Table IV. The overall classification accuracy is increased. To the CVD domain, the accuracy of walking increases to 100 % from 86.5 % when the RT and µD features were applied to fusion. Compared with using RT and µD feature independently, the fusion with CVD also improves the accuracy of standing up (by 38 % and 11 %), sitting down (by 38 % and 5 %), picking up an object (by 20 % and 10 %) and drinking (by 14 % and 4 %) for both RT and µD, respectively. The fusion of CVD with the other two domains could cover the deficiencies of each other. However, when the fusion was applied to features of the µD and RT, it exacerbates the poor accuracy. This exacerbation might be caused by features with similar drawbacks.

Table IV Comparison of accuracy for individual activities and averaged across the activities using each data domain independently, with feature fusion, and with customised hierarchical classification.

Accuracy (%)	Walking	Standing up	Sitting down	Picking up an object	Drinking	Fall	Overall Performance
µD	**100**	82.2	88.4	63.6	79.7	91.9	**84.3**
CVD	86.5	**92.7**	**89.0**	**63.7**	**80.7**	**92.8**	84.2
RT	**100**	53.1	51.4	53.3	60.9	87.1	67.6
µD+CVD	**100**	**92.7**	**92.9**	**74.2**	**83.8**	**97.1**	**90.1**
µD+RT	**100**	51.5	28.6	42.7	58.1	86.2	61.2
CVD+RT	**100**	90.9	89.3	73.5	75.2	92.8	87.0
µD+CVD+RT	**100**	94.2	91.7	**77.3**	**84.9**	97.1	90.9
Custom hierarchy	**100**	**95.5**	**95.2**	76.9	84.6	**100**	**92.0**

** The accuracy of three domains are recorded from SVM classifiers with Chi-Square algorithm, due to their better performance compared to others.

3.4 Hierarchical structure

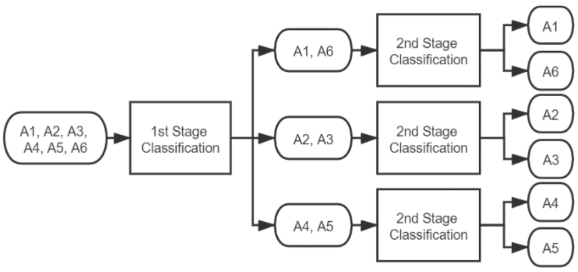

Fig. 6 A custom hierarchical classification structure based on the result and analysis.

Based on the previous results, a hierarchical structure is introduced to optimise the overall performance. As the number of features increases, the same feature pool or classifier will not be the optimum for classifying all the activities. Unlike the conventional 'flat' classification, which classifies each example to all its available labels at the same time, the hierarchical classification uses data taxonomy to create a hierarchy of classifiers [22]. In this paper, the hierarchical structure firstly separates activities into several subgroups, and then each subgroup can implement distinct features and classifiers, which are more suitable for the subgroup, to improve the overall performance. Fig. 6 demonstrates the structure of hierarchical classification. The activities were divided into three subgroups (A1, A6; A2, A3 & A4, A5) at the first stage due to their similarity and false alarm rate. The classifiers of all stages were SVM, with CS algorithm used for feature selection at each classification stage. The selections of each stage were independent, which means the discarded features at the first stage would come back to the feature pool for selection at the second stage. The fusion data of µD + CVD + RT was used for the feature sets at the first stage, and at the second stage for A2 and A3, A4 and A5. At the upper second stage, which classifies the A1 and A6, the feature fusion of CVD + µD was employed with less computation burden due to the smaller number of features set. Fig. 7 presents the result of the hierarchical classification. The overall performance of classification is increased to approximately 92 %, which improves by ~1 % the result of the best fusion combination. It is observed that the accuracies of all the activities were increased except A4 and A5. The majority of misclassification of A4 and A5 was generated at the second stage. It was observed from Fig. 1-3 that the RT, µD and CVD of A4 and A5 were similar. This situation became more serious when the participants were the elders. From Fig. 4, it was obvious that the spectrograms of A4 and A5 were almost the same, which means the differences between features extracted from them were little, increasing the possibility of misclassification. Besides, misclassification also happened at the first stage between subgroups (A2, A3 &A4, A5). If an activity was misclassified at the first stage, it would also be an error at the second stage, where the errors were accumulated.

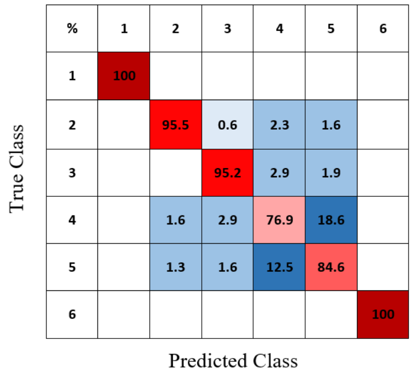

Fig. 7 Confusion matrix of hierarchical classification.

IV Conclusion and Future work

In this paper, we presented and compared different feature extraction, selection and classification approaches for discriminating daily human movements. The preliminary results show that RT information is sensitive to the moving target, which means translating in space across range bins, whereas it cannot classify accurately in-place body motions. In contrast, CVD is more sensitive to body movements in place, but weaker in dealing with moving target detection. Spectrogram can not only detect moving targets, but also observe body movements. Additionally, fusion approach and hierarchical classification were also implemented. Improvements in accuracy, when the feature-level fusion was applied, were between ~6 and 23 %, whereas the accuracy of the best fusion combination was 90.9 %, overcoming the limitation in using single domain approach. When the hierarchical classification was implemented, the accuracy was in the order of approximately 92 %. Further work will focus on more domains, approaches of selection, classification, data fusion, and how the possible combination of these factors could further optimise the accuracy.

Acknowledgment

The authors would like to thank the British Council 515095884 and Campus France 44764WK – PHC Alliance France-UK for their financial support.

References

[1] S. C. Mukhopadhyay, "Wearable Sensors for Human Activity Monitoring: A Review," in *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321-1330, March 2015.

[2] L. Liu, F. Hu and J. Zhao, "Action Recognition Based on Features Fusion and 3D Convolutional Neural Networks," *2016 9th International Symposium on Computational Intelligence and Design (ISCID)*, Hangzhou, 2016, pp. 178-181.

[3] V. C. Chen, F. Li, S. -. Ho and H. Wechsler, "Micro-Doppler effect in radar: phenomenon, model, and simulation study," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 42, no. 1, pp. 2-21, Jan. 2006.

[4] Z. Sun, J. Wang, Y. Zhang, J. Sun, C. Yuan, Y. Bi, "Multiple walking human recognition based on radar micro-Doppler signatures," in *Sci. China Inf. Sci*, vol. 58, no. 12, pp. 1-13, Dec. 2015.

[5] Y. He, X. Li, X. Jing, "A Mutiscale Residual Attention Network for Multitask Learning of Human Activity Using Radar Micro-Doppler Signatures," in *Remote Sens*, vol. 11, pp. 2584, 2019.

[6] C. Clemente, A. Balleri, K. Woodbridge, and J. J. Soraghan, "Developments in target micro-doppler signatures analysis: radar imaging, ultrasound and through-the-wall radar," in *EURASIP Journal on Advances in Signal Processing*, 2013.

[7] X. Li, Y. He, X. Jing, "A Survey of Deep Learning-Based Human Activity Recognition in Radar", in *Remote Sens*, vol. 11, pp. 1068, 2019.

[8] Y. Lin, J. Le Kernec, S. Yang, F. Fioranelli, O. Romain and Z. Zhao, "Human Activity Classification With Radar: Optimisation and Noise Robustness With Iterative Convolutional Neural Networks Followed With Random Forests," in *IEEE Sensors Journal*, vol. 18, no. 23, pp. 9669-9681, 1 Dec.1, 2018.

[9] F. Fioranelli, M. Ritchie and H. Griffiths, "Performance Analysis of Centroid and SVD Features for Personnel Recognition Using Multistatic Micro-Doppler," in *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 5, pp. 725-729, May 2016.

[10] B. Jokanović and M. Amin, "Fall Detection Using Deep Learning in Range-Doppler Radars," in *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 180-189, Feb. 2018.

[11] H. Du, T. Jin, M. Li, Y. Song and Y. Dai, "Detection of multi-people micro-motions based on range–velocity–time points," in *Electronics Letters*, vol. 55, no. 23, pp. 1247-1249, 14 11 2019.

[12] A. Seifert, A. M. Zoubir and M. G. Amin, "Radar-based human gait recognition in cane-assisted walks," *2017 IEEE Radar Conference (RadarConf)*, Seattle, WA, 2017, pp. 1428-1433.

[13] S. Bjorklund, H. Petersson, and G. Hendeby, "Features for micro-Doppler based activity classification," *in IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1181–1187, 2015.

[14] H. Li, A. Shrestha, H. Heidari, J. Le Kernec and F. Fioranelli, "Magnetic and Radar Sensing for Multimodal Remote Health Monitoring," in *IEEE Sensors Journal*, vol. 19, no. 20, pp. 8979-8989, 15 Oct.15, 2019.

[15] J. Le Kernec *et al.*, "Radar Signal Processing for Sensing in Assisted Living: The Challenges Associated With Real-Time Implementation of Emerging Algorithms," in *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 29-41, July 2019.

[16] S. Z. Gürbüz, B. Erol, B. Çağlıyan and B. Tekeli, "Operational assessment and adaptive selection of micro-Doppler features," in *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1196-1204, 12 2015.

[17] H. Li, A. Shrestha, H. Heidari, J. L. Kernec and F. Fioranelli, "A Multisensory Approach for Remote Health Monitoring of Older People," in *IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology*, vol. 2, no. 2, pp. 102-108, June 2018.

[18] M. A. A. H. Khan *et al.*, "RAM: Radar-based activity monitor," *IEEE INFOCOM 2016 - The 35th Annual IEEE International Conference on Computer Communications*, San Francisco, CA, 2016, pp. 1-9.

[19] S. Visalakshi and V. Radha, "A literature review of feature selection techniques and applications: Review of feature selection in data mining," *2014 IEEE International Conference on Computational Intelligence and Computing Research*, Coimbatore, 2014, pp. 1-6.

[20] F. Castanedo, "A review of data fusion techniques," in *The Scientific World Journal*, 2013.

[21] R. C. Luo, Chih-Chen Yih and Kuo Lan Su, "Multisensor fusion and integration: approaches, applications, and future research directions," in *IEEE Sensors Journal*, vol. 2, no. 2, pp. 107-119, April 2002.

[22] C. N. Silla, A. A. Freitas, "A survey of hierarchical classification across different application domains, " in *Data Mining and Knowledge Discovery,* vol. 22, pp. 31–72, 2011.