Hierarchical Classification on Multimodal Sensing for Human Activity Recogintion and Fall Detection

Haobo Li, Aman Shrestha, Francesco Fioranelli, Julien Le Kernec, Hadi Heidari
School of Engineering, University of Glasgow, Glasgow, United Kingdom
francesco.fioranelli@glasgow.ac.uk

Abstract— **This paper presents initial results on the usage of hierarchical classification for human activities discrimination and fall detection in the context of assisted living. Multimodal sensing is proposed by combining data from a wearable device and a radar system. The effect of different approaches in selecting the activities in each sub-group of the hierarchy are explored and reported as preliminary results in this work, while a more detailed investigation is undergoing. 1.2-2.2% improvement in accuracy with SVM and DL classifiers compared with the conventional case of activity classification is reported; subsequent improvement (1.6%) occurs when using SVM-SFS in the second stage of hierarchical classification.**

Keywords—multi-modal sensing, hierarchical classification, human activity recognition, fall detection, machine learning.

I. INTRODUCTION

The ratio of people over 65 years old in the UK to the national population is currently 1 to 6, and is expected to reach 1 to 4 by 2050 [1]. In this context where self-protection awareness is lacking as well as timely medical assistance, more than 50,000 older people are sent to the hospital through emergency services annually after fall accidents [2]. Serious physical trauma, such as brain concussions [2, 4] and hip fractures [3-4], can result from falls, as well as psychological trauma and reduced motivation to exercising and performing rehabilitation. This can lead to reduced life expectancy, with about 58% of elderly people over 80 years reported to die after serious falls [4].

Therefore, an effective and reliable automatic fall detection [5] system is in high demand to significantly reduce the consequences after an incident by shortening the response time of personal nurses or ambulances services [5]. Besides fall detection, the system proposed also intends to characterize and track normal daily behavior patterns for the monitored elderly people at high risk of falling. Human activity recognition (HAR) can provide fine-grained information concerning their well-being and early warning of anomalies, especially when they live alone. This can in turn enable personalized diagnostic and treatment, by identifying specifying issues affecting the person as highlighted by changing in their normal behavior.

It has been proved in our previous work that using multimodal sensing can maximize the classification performance and overcome the drawbacks of each sensing technologies [6]. Inertial Measurement Units (IMU), [7] the most representative wearable sensor-based device, is easy to be miniaturized and integrated into daily electronic products (e.g. smart phone or wristband, or even smart clothes); however, such wearable devices require to be worn (e.g. wrist, ankle, thigh or foot), and therefore the end-user needs to remember to wear or take the sensor with them. Different from wearables and camera-based systems (e.g. Microsoft Kinect), Frequency Modulated Continuous Wave (FMCW) radar sensors collect range and speed information from echo signals returned by specific targets. This sensing modality is contactless, does not require end users' compliance, and at the same time avoids the image generation of photos and videos, which may raise privacy issues [8].

Hence, in this paper we choose to validate a multimodal sensing strategy combining wearable devices and radar sensors for activity recognition and fall detection. A hierarchical classification model is introduced to approach the problem by dividing the different human activities into several sub-groups. This is expected to improve the overall classification performance by combining different sets of features and sensors for each sub-group of activities, and is a further step compared with our analysis in [6, 9]. To be closer to the real world usage, the classifier is tested with data from one of the participants out of 20, and trained with the remaining data. Furthermore, for the purpose of simultaneously improving classification performance and reducing the computational intensity, Sequential Forward Selection (SFS) [10] is chosen as the feature selection method to generate the best feature combinations for each sub-group of activities.

The rest of this paper is organized as follows. Section II describes briefly the experimental setup and the features used. Section III presents the hierarchical classification approach and discusses some preliminary results. Section IV draws conclusions and outlines future work.

II. EXPERIMENTAL SETUP AND FEATURE EXTRACTION

Ten daily activities are considered in this paper: walk (A1), walk while carrying an object (A2), sitting (A3), standing (A4), pick up an object (A5), tie shoelaces (A6), drinking water (A7), answer a phone call (A8); fall (A9); crouch and stand back up (A10). They are performed by 20 male participants aged from 21 to 35 and repeated three times. In order to produce a more challenging classification task, some of the activities are designed to be similar in pairs, for instance A3, A5, A6 and A9 all involve the rapid vertical motion of the torso downwards, which makes them harder to distinguish from one another. A more detailed description of the activities can be found in [11].

Data is simultaneously collected by different sensing devices, in particular, a commercial high performance IMU from x-IO technology, and a FMCW radar sensor working at 5.8 GHz with 400 MHz bandwidth and 1 ms chirp duration. The wearable device is comprised of a tri-axial accelerometer, gyroscope and magnetometer, which the sampling frequency is set as 50 Hz. The experiment took place at the University of Glasgow Communication, Sensing and Imaging (CSI) Laboratory. An overview of the experimental environment is shown in Fig. 1, where the left-hand figure shows the radar FMCW: blue box and its Yagi-antennas, and the right-hand side figure shows one of the participants with wearable device on the wrist of the dominant hand.

The data contains 10 degrees of freedom (DOF), notably, three directions (X, Y and Z) for each sensors in IMU (3x3), plus the radar system, whereas the number of observations are equal to 600 (20 participants x 10 activities x 3 repetitions).

Fig. 1. Radar sensor and antennas (left) and participant with wearable IMU device at wrist (right)

Prior to the hierarchical classification, features are derived from the raw data in terms of numerical values. There are 64 features corresponding to accelerometer, gyroscope and magnetometer respectively, while the radar has 24 features. Details of the features extracted are available in [10]. Fusion takes place at feature level to combine the features from individual sensors to a feature matrix.

III. HIERARCHICAL CLASSIFICATION

Unlike conventional classification methods illustrated in the top of Fig. 2, which inputs all the activities to the classifier at the same time, 'hierarchical' approach means that firstly the activities are divided into several sub-groups according to the similarity between activities or misclassification, for instance, three sub-groups in Fig. 2, which is known as the 1^{st} classification stage. After this, different classifiers are implemented to predict the labels inside each sub-groups (2^{nd} stage). A single feature set and one specific classifier are used in the normal classification method; however, as the number of 'confusing' classes increase, this same feature set or classifier will not be necessarily optimal for classifying all the activities. Hierarchical model enables to utilize different feature sets and classifiers for different classes being considered, trying to increase the overall performance. Furthermore, SFS could be applied to find optimal feature combinations regarding to activities in each sub-groups, instead of using the same features for all the activities. In fact, the set of optimal features can be different for each sub-group of activities, rather than using the same, fixed one, for all of them. During the first and second classification stages, a robust Quadratic-kernel SVM algorithm [11], a K-Nearest Neighbor (KNN) method [12] with K=5 and a linear discriminant analysis classifier (DL) [12] are adopted to train the model and predict the labels of tested data accordingly.

In this paper, the validation of classification performances is made with a "Leave one subject out" approach [13]. We use data from one unknown participant to evaluate the robustness of classifiers trained with data from other the participants, as it would happen operationally, because it is unfeasible for classifiers to be trained with data from all specific end users. The validation accuracy in this section is based on the average value of 20 iterations, in which each iteration corresponds to testing the 'Leave one person out' method on one of the participants, until everyone has been tested.

Additionally, the way the 10 activities are divided into sub-groups will affect the performance. In this paper, we have tried four different approaches to divide the activities into a maximum of 5 sub-groups, as reported in Table I. Grouping is based on two principles. Firstly, the similarity between activities so that highly similar activities are divided into one group, such as No.1. Then, the misclassifications between activities when no hierarchy is used, so that these activities are assigned to the same group, such as No.2, 3 and 4 in order to optimize the feature selection for improved performance.

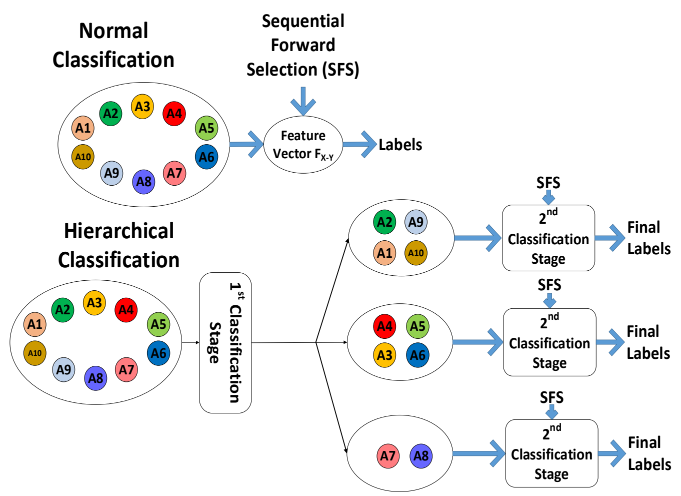

Fig. 2. Representation of conventional and hierarchical classification

TABLE I ACTIVITIES DIVISION METHODS

Activities division methods	Activities in Sub-group 1	Activities in Sub-group 2	Activities in Sub-group 3	Activities in Sub-group 4	Activities in Sub-group 5
No.1	1&2	3&4	5&6	7&8	9&10
No.2	1&2&10	3&5&6&9	4&7&8	N/A	N/A
No.3	1&2	3&4&5	6&7&8	9&10	N/A
No.4	1&2&9&10	3&4&5&6	7&8	N/A	N/A

Fig. 3 presents the classification results for the different activity sub-groups (Table I) and different classifiers. It is observed that Support Vector Machine (SVM) outperforms the other two classifiers. Furthermore, if the activities are divided as in case No.4 of Table I (indicated by V in Fig. 3), the best classification performance is obtained, with 1.3%

(SVM), 0.5% (KNN), and 2.4% (DL) enhancement compared to the conventional classification with no hierarchy (indicated by I in Fig. 3). Activities division case No.1 yielded the lowest accuracy due to poor performance of the first classification stage.

The best activity division case (No.4 in Table I), is chosen to apply SFS on the feature sets used for each classification at the second stage. In Fig. , we compare the results when using SFS for a normal classification approach, and for the hierarchical approach, with activities divided according to the 4th method in Table I. The results show that the improvement due to hierarchical approach are dependent on the type of classifier chosen, with the best improvement obtained for SVM (approximately 2%). If we compare this with the results obtained for no SFS and no hierarchical approach (see case I in Fig. 3), the overall classification accuracy is improved by about 2.7% (SVM), 0.8% (KNN) and 3.4% (DL). If we compare the SVM result in Fig. 4 (for both SFS and hierarchical approach) with the results in Fig. 3 (with hierarchical approach applied but not SFS), we can see a subsequent improvement. This appears to suggest that SFS using a robust classifier (e.g. SVM) is more effective when applied in conjunction with hierarchical approach; additional work is undergoing to verify this preliminary result.

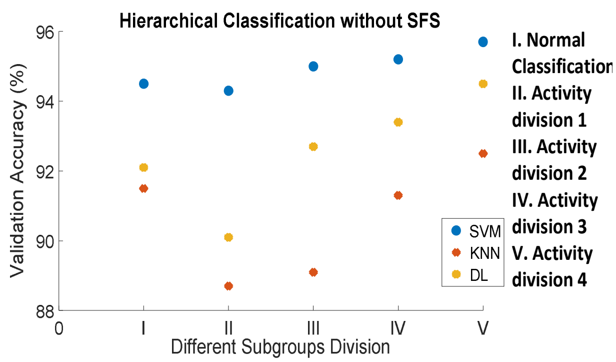

Fig. 3. Hierarchical classification with different sub-groups and classifiers

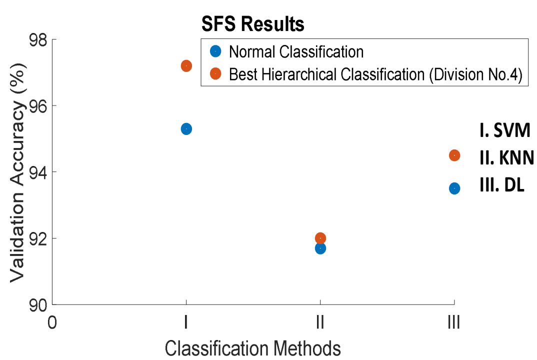

Fig. 4. SFS for best hierarchical division through different classifiers

It is interesting to find out which sensor contributes the most to the hierarchical classification. From Table II, four existing sensors (accelerometer, gyroscope, magnetometer and radar) are used to evaluate the classification performance individually, as well as considering the combination of three inertial sensors and inertial sensors plus radar. From accelerometer and magnetometer, approximately 2-3% improvement after using hierarchical classification is reported, whereas the accuracy of gyroscope and radar decreases. However, the usage of three inertial sensors and all sensors perform better with the help of hierarchical approach and feature selection.

The confusion matrix of the best-case scenario is illustrated below in Fig. 5, the row elements are the output class while the columns represent the target class. The main misclassification occurs between distinguishing A3, A4 and A5, as well as A7 and A8. However, the most critical action - fall detection (A9) - yielded high sensitivity (98.3) and low false alarms (1.6).

TABLE II BEST HIERARCHICAL AND NORMAL CLASSIFICATION WITH DIFFERENT SENSORS

Accuracy [%]	Acce	Gyro	Magn	Radar	IMU	All
Normal Classification	86.7	80.7	79.8	88.5	88.8	94.5
Best Hierarchical after SFS	89.7	78.33	82	87	91.5	97.2

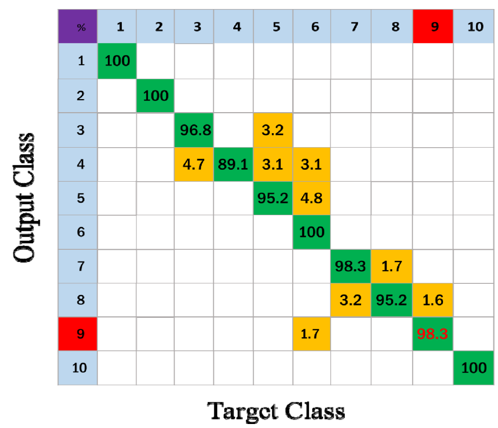

Fig. 5. Confusion matrix of best hierarchical classification with SFS

IV. CONCLUSIONS AND FURTHER WORK

In this paper we presented some initial results on using hierarchical classification for activity recognition for assisted living. Data from a wearable sensor and a radar system have been combined in multimodal sensing approach. The effect of different schemes to divide the activities in sub-groups and the usage of SFS for feature selection have been explored and commented upon. Although the improvements in terms of overall accuracy are modest (about 2-3% at most), it is interesting to explore whether this hierarchical approach could yield better improvements in different scenarios, perhaps starting from a baseline accuracy much lower than 90% in a more challenging scenario. As the number of possible activities to identify and the number of available sensors increase, the usage of hierarchical approaches is expected to provide additional tools to enhance performance.

Future work will investigate in more details how the sampling rates, possible combination of sensors, selection of features, and the division into sub-groups of activities can be further optimized to improve performance. Additional data will be collected using a variety of radar sensors and multiple wearables on different body parts.

References

[1] Office of National Statistics, "Overview of the UK population: July 2017," 2017.

[2] M. Terroso, N. Rosa, A. T. Marques, and R. Simoes, "Physical consequences of falls in the elderly: a literature review from 1995 to 2010," European Review of Aging and Physical Activity, vol. 11, no. 1, p. 51, 2014

[3] J. A. Stevens, M. F. Ballesteros, K. A. Mack, R. A. Rudd, E. DeCaro, and G. Adler, "Gender differences in seeking care for falls in the aged Medicare population," *American journal of preventive medicine,* vol. 43, no. 1, pp. 59-62, 2012.

[4] C. S. Florence, G. Bergen, A. Atherly, E. Burns, J. Stevens, and C. Drake, "Medical costs of fatal and nonfatal falls in older adults," *Journal of the American Geriatrics Society,* 2018.

[5] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of neuroengineering and rehabilitation,* vol. 9, no. 1, p. 21, 2012.

[6] H. Li *et al*., "Multisensor data fusion for human activities classification and fall detection," *2017 IEEE SENSORS*, Glasgow, 2017, pp. 1-3.

[7] H. Heidari, E. Bonizzoni, U. Gatti and F. Maloberti, "A CMOS Current-Mode Magnetic Hall Sensor With Integrated Front-End," in *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 62, no. 5, pp. 1270-1278, May 2015.

[8] E. Cippitelli, F. Fioranelli, E. Gambi and S. Spinsante, "Radar and RGB-Depth Sensors for Fall Detection: A Review," *IEEE Sensors Journal*, vol. 17, no. 12, pp. 3585-3604, 2017.

[9] 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.

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

[11] T. D. Bufler and R. M. Narayanan, "Radar classification of indoor targets using support vector machines," *IET Radar, Sonar & Navigation,* vol. 10, no. 8, pp. 1468-1476, 2016.

[12] A. Shrestha, J. Le Kernec, F. Fioranelli, E. Cippitellii, E. Gambi and S. Spinsante, "Feature diversity for fall detection and human indoor activities classification using radar systems," *International Conference on Radar Systems (Radar 2017)*, Belfast, 2017, pp. 1-6.

[13] K. Altun and B. Barshan, "Human activity recognition using inertial/magnetic sensor units," in *International Workshop on Human Behavior Understanding*, 2010, pp. 38-51: Springer.

.