

Model-based and Class-based Fusion of Multisensor Data

Athina Tsanou¹, Angelos Chatzimichail¹, Georgios Meditskos¹, Stefanos Vrochidis¹, Ioannis Kompatsiaris¹

¹Information Technologies Institute, Center for Research and Technology Hellas

Abstract. In the recent years, the advancement of technology, the constantly aging population and the developments in medicine have resulted in the creation of numerous ambient assisted living systems. Most of these systems consist of a variety of sensors that provide information about the health condition of patients, their activities and also create alerts in case of harmful events. Successfully combining and utilizing all the multimodal information is an important research topic. The current paper compares model-based and class-based fusion, in order to recognize activities by combining data from multiple sensors or sensors of different body placements.

Keywords: Activity recognition · Wearable sensors · Fusion.

1 Introduction

Population of elderly people is constantly rising and will continue to increase significantly according to expectations. Although this creates a lot of problems in the health care, including higher costs and larger number of people with difficulties in self-serving, the advancement in technology is able to provide solutions [1]. Systems that employ artificial intelligence applications, have the ability to recognize the daily activities of a subject, provide information about heart rates, blood pressure or temperature and detect the occurrence of a harmful event, such as fall.

Assisted living systems utilize a variety of sensors, devices and technologies. Inertial sensors are usually included in such systems, however they can also be used separately to provide information about the activities performed by a person. Nowadays, such sensors can be easily found in mobile and wearable devices. Most commonly used sensors are accelerometers, gyroscopes and magnetometers. Accelerometers, which are the most popular, are quite effective in recognizing activities with repetitive body movement [2] while gyroscopes are capable of recognizing the orientation of an object [3]. Accelerometers and gyroscopes can be used separately and produce adequate results, while magnetometers are known to perform poorly when used individually [4]. Inertial sensors in general, are very prone to environmental noise, thus combining them tends to increase the accuracy of the recognition rate. Combination of sensors is algorithmically achieved through combination of the features extracted from each sensor, referred to as early fusion, or combination of the classification results of each individual sensor, also known as late fusion. It is easily understood that late fusion can be also applied to an individual sensor, combining the results of different classification algorithms and improve its performance.

In human activity recognition problems, fusion can also be applied to combine the results of the same sensor placed on different body parts, since there is claim that the placement of accelerometer affects its performance [5]. Although it is not considered mistaken to ignore the location of the sensor when analyzing its recordings, there are studies that investigate the affect the body placement has on the recognition of activities. In [5], the authors propose a late fusion methodology that combines accelerometers placed on ankle, wrist and chest. Authors in [4] investigate how five different body locations of a smartphone affect the performance of its built-in sensors in recognizing particular activities. Performance of individual sensors and their combination by concatenating feature vectors is also studied. [6] and [7] combine accelerometers of different body placements by applying early fusion.

In the current study, late fusion is employed to combine three inertial sensors, namely accelerometer, gyroscope and magnetometer, by body placement. Furthermore, the fusion of different placements of the same sensor is investigated. For the late fusion, two weighting schemes are incorporated and compared with

a baseline late fusion method. The first one is a model-based method, the weighted accuracy [1], which reflects the total performance of a classifier. The second scheme is a class-based method, recently proposed by the current authors in [16], that utilizes the detection rate of a class, so as to emphasize the ability of a classifier in detecting certain classes and not its overall performance. In the latter, we suggested using fusion weights that are equal to the supplement of the class detection rate. We incorporated the suggested weights in a typical late fusion function and in a framework with posterior adapted class probabilities. The suggested method was compared with known late fusion methods, like averaging and stacking and with a novel framework suggested in [5]. Finally, we compare the model and class based fusion schemes with a simple form of late fusion, the averaging of class probability vectors produced by different models.

The rest of the paper is organized as follows: Section 2 includes a brief mention of related work. Section 3 explains the theory used and Section 4 presents the results of the application. Section 5 concludes the paper.

2 Related work

Although fusion has recently started gaining popularity, there is already a wide variety of available research work. As already stated, fusion can be applied to combine different sensors, even quite heterogeneous ones, sensors placed on different locations or improve the performance of an individual sensor by combining results of several algorithms. A methodology for recognizing activities from wearable sensors is proposed in [8]. The methodology is based on the late fusion of classification results obtained from Neural Networks (NN) and Hidden Markov Models (HMM) using two sensors placed on different body locations of the participants. In [9] a wireless sensor network is developed to observe the status of persons in need of assistance. Accelerometer data are combined with images from three different cameras to detect falls and improve the system's ability to create true alerts. In [10] data from a wearable inertial sensor and two cameras are combined in order to detect events. For the fusion step a probabilistic scheme is proposed by the authors. Many studies use audio visual sensors to recognize activities, although this process is usually more time consuming and the equipment needed requires bigger budgets. As already mentioned, input from wearable sensors is often combined with video images, especially in platforms that provide support at home for people with disabilities. In [11] the authors implement three types of fusion, early, intermediate and late, to analyze input from wearable cameras and recognize activities. HMM algorithm is incorporated in the process to classify the data. Different forms of fusion allow for the combination of multiple modalities in different levels. In [13] accelerometer and video data are combined in different stages of the classification process. The authors combine the different inputs before the feature extraction step, after feature extraction by concatenating the different vectors and at results' level by combining the algorithm outputs.

In [4] the authors explored the individual performance of each one of the three smartphone sensors, accelerometer, gyroscope and magnetometer, the combination of two of them, and the affect of the location of the sensor in the recognition process. For the combination of sensors or placements, the authors use the simplest form of early fusion, the concatenation. The current paper, on the contrary, uses late fusion to combine three sensors rather than two. In [12], the authors use built-in sensors of smart devices, separately and combined, and apply several well known classifiers to recognize activities. With the assistance of GPS and light sensors the activities can be further categorized to indoor or outdoor. Accelerometer, gyroscope and magnetometer data are combined with light and pressure data, and after being interpolated and filtered, features are extracted and combined with GPS information in order to enter a classification algorithm. The authors don't apply any fusion techniques, they focus however on the preprocessing of the data so as to eliminate noise and heterogeneity. Accelerometers and gyroscopes are most often combined in fusion schemes, probably due to their satisfactory performance in the daily activities recognition. Early and late fusion is applied in [17] to combine accelerometer and gyroscope features. The authors use concatenation for early fusion, a weighted scheme for late fusion and a descriptor-based framework for the activity recognition. Authors in [18] fuse accelerometer and gyroscopes to recognize activities and detect falls and focus on the importance of the window size for signal segmentation.

Weights are quite often utilized in fusion schemes, especially in late fusion ones. Out of bag errors acquired from the random forest algorithm are used in order to combine classifier results of different modalities in [14]. The classification problem is not relevant to the activity recognition, the methodology though could be easily implemented on multisensor data. In [1] the authors make use of a multisensor platform and apply several fusion weights combined with different fusion functions in order to recognize activities. The authors also propose a genetic algorithm to calculate weights. The weighted accuracy included in the latter, is also utilized in the current paper.

3 Methodology

The current human activity recognition framework comprises of the following steps:

1. Sliding window segments
2. Feature extraction
3. Classification algorithm
4. Model-based or class-based fusion

Sliding windows of 2 seconds without overlap were taken in order to extract features, similar to [5]. Time domain features, mentioned in Table 1, were extracted from the sliding windows without any further filtering or preprocessing of the data. The initial dataset was then segmented in the required subsets, responding to the sensor and body placement we wanted to analyze. Several multilabel classification algorithms were tested and the four that achieved better results are included here: Support Vector Machines (SVM), Random Forests (RF), C5 trees and k-Nearest Neighbors (kNN). Each classifier was applied separately to a sensor and the classification results of the algorithms were combined afterwards. For the fusion step, two types of fusion were tested. **Model-based**, which characterizes the overall performance of the classifier and **class-based** which pays attention to the recognition of specific classes. The results were compared with those of the simple late fusion method of averaging class probabilities. To derive both types of weights for the fusion step, the typical steps of a classification framework were applied. An algorithm was trained on the trainset and then applied to the testset in order to predict the types of activities. The fusion schemes applied, combine sensors according to the following two scenarios:

1. Different sensors with the same placement
2. Identical sensors of different placement

Model-based fusion For the model-based approach the weighted accuracy was used. The accuracy (Eq. 2) of a classifier applied to a sensor, divided by the sum of accuracies, was multiplied by the class probability vectors (Eq. 1) and the products of all three sensors were finally added together to create a final class probability vector [1]. The class with the maximum probability was assigned to each test case. This method gives advantage to the model that has the best performance overall. The formula for weighted accuracy, as described in [1] is given in Eq. 3 and is calculated for each one of the i models:

$$P_{ij} = \{p_{i1}(x_1), \dots, p_{ik}(x_n)\}, i = 1, \dots, m \quad (1)$$

Table 1. Extracted Features

Features
Maximum
Mean
Median
Minimum
Standard Deviation
Variance

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (2)$$

$$WACC = \frac{Accuracy^{(i)}}{\sum_{i=1}^m Accuracy^{(i)}} \quad (3)$$

Class-based fusion For the class-based fusion, a novel method proposed by the authors of the current paper in [16] is applied. Class-based methods pay attention to the recognition of each class, which is usually characterized by F1score [5] or balanced accuracy. This method multiplies the class probabilities with the supplement of the detection rate, which was chosen as weight, to assist the recognition of classes not so easily predicted. This is performed for each sensor separately and the weighted probability vectors of all sensors are afterwards summed together. Again, to assign a class to a test case, we find the class with the maximum fused probability. The detection rate is defined in Eq. 4.

$$DR = TP/(TP + TN + FP + FN) \quad (4)$$

The weights are calculated for each class j by calculating the supplement of the class detection rate (Eq. 5).

$$W_{ij} = 1 - DR_{ij} \quad (5)$$

Both types of weights are then multiplied with the class probabilities of an algorithm (Eq. 6). In the model based fusion, the class probability vectors are multiplied by the same number, the WACC, while in the class-based weight, each class probability vector is multiplied by a different weight.

$$P_w = W_{ij}P_{ij} \quad (6)$$

Averaging Each model of those that will be combined, produces k ($j = 1, \dots, k$) class probability vectors (Eq. 1). To combine the results, a probability vector is calculated for each class, by averaging the respective class probability vectors of the m models combined (Eq. 7).

$$P_j = \frac{\sum_{i=1}^m P_{ij}}{m} \quad (7)$$

To evaluate the performance of the individual classifiers and the fusion results, we report two typical evaluation metrics. The accuracy of the model (Eq. 2) and the F1-score (Eq. 8). For multilabel classification, the F1-score is calculated for all classes and then averaged. The F1-score embodies both sensitivity and specificity and especially in multiclass problems, is considered more indicative of the accuracy metric, since it assesses the recognition of each class. In multiclass problems, high accuracy values may arise, while few of the classes may have not been recognized at all, a finding that is also confirmed in the current paper, as it can be seen in the following section.

$$F1 - score = 2 * precision * recall / (precision + recall) \quad (8)$$

4 Application

In order to compare the model-based fusion method of weighted accuracy with the class-based fusion method, proposed by the authors in [16], and with the simple late fusion method of averaging and assess the influence of location on the sensors' performances, we utilize the MHEALTH dataset [15]. The MHEALTH contains recordings of numerous sensors, some of which were placed at different parts of the subject's body. The activities were performed by ten participants, nine of which have constituted the training set and the

recordings of one participant constituted the test set. The subjects performed 13 daily activities, 8 of which are kept as a subset for the current application. The activities were "Standing still", "Sitting and relaxing", "Lying down", "Walking", "Climbing stairs", "Cycling", "Jogging" and "Running". The time domain features that are mentioned in Table 1 are extracted from a sliding window of 2 seconds without overlap, which in the current datasets responds to 100 recordings, since the sampling frequency is 50 Hz.

Since it has been reported that body placement of a sensor affects its performance [5], we chose to combine a) results of accelerometers, gyroscopes and magnetometers of the same body location and b) accelerometers placed on three different locations, as well as gyroscopes placed at two different locations. More specifically, the results reported here for a) refer to the fusion of

1. The three aforementioned sensors placed on the left ankle
2. The three sensors placed on the right lower arm and the results reported for b) refer to

the fusion of

1. Accelerometer placed on chest, with that placed on the left ankle and the one of the right lower arm
2. Gyroscope placed on left ankle with the gyroscope of the right lower arm.

The flowchart in Fig. 1 provides a graphic demonstration of the procedure. The flowchart describes the fusion of two sensors as an example, while three sensors are fused in the current application. The same procedure is repeated for the sensors on the right lower arm and for the fusion of identical sensors of different locations.

Table 2 presents the evaluation metrics of the four classifiers applied on each sensor of the left ankle separately and of the three fusion methods, the weighted accuracy, which characterizes the performance of the algorithm, the detection rate based weighted fusion which evaluates the ability of an algorithm to detect each class and averaging. Table 3 contains the respective results for the sensors placed on the right lower arm. In most of the classifiers, fusion schemes exceed the best performing classifier on individual sensor data. In half of the cases the class-based fusion outperforms the model-based one.

Table 2. Fusion of sensors placed on left ankle

	SVM		RF		C5		KNN	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Accelerometer	0.5750	0.5004	0.775	0.7670	0.5667	0.5649	0.7500	0.7489
Gyroscope	0.7000	0.6900	1.0000	1.0000	0.9542	0.9537	0.8583	0.8167
Magnetometer	0.3458	0.2724	0.825	0.8245	0.8208	0.8166	0.7708	0.7706
WACC	0.7542	0.7311	0.9875	0.9874	0.9708	0.9706	0.8792	0.8486
DR	0.7667	0.7474	0.9500	0.9500	0.8958	0.8834	0.8833	0.8535
Averaging	0.7500	0.7272	0.9583	0.9585	0.9167	0.9112	0.8875	0.8638

Table 3. Fusion of sensors placed on right lower arm

	SVM		RF		C5		KNN	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Accelerometer	0.6417	0.5582	0.8750	0.8333	0.8750	0.8333	0.9000	0.8845
Gyroscope	0.5292	0.4200	0.8500	0.8085	0.8375	0.7978	0.7083	0.6823
Magnetometer	0.3542	0.2940	0.6458	0.6257	0.6167	0.5979	0.5167	0.4812
WACC	0.6583	0.5763	0.8750	0.8333	0.8958	0.8739	0.9542	0.9529
DR	0.6875	0.6081	0.8750	0.8322	0.8875	0.8637	0.9208	0.9174
Averaging	0.6875	0.6057	0.8750	0.8322	0.8833	0.8595	0.8833	0.8765

SVM has the worst performance among all four classifiers at all cases. In general, although the accuracy and F1-score may indicate good performance, some classes were not recognized at all from certain sensors and classifiers. Three classifiers, namely SVM, RF and C5, applied on the accelerometer placed on the left ankle, failed to recognize one activity, "sitting and relaxing", while SVM was not able to predict at all two more activities: "climbing stairs" and "cycling".

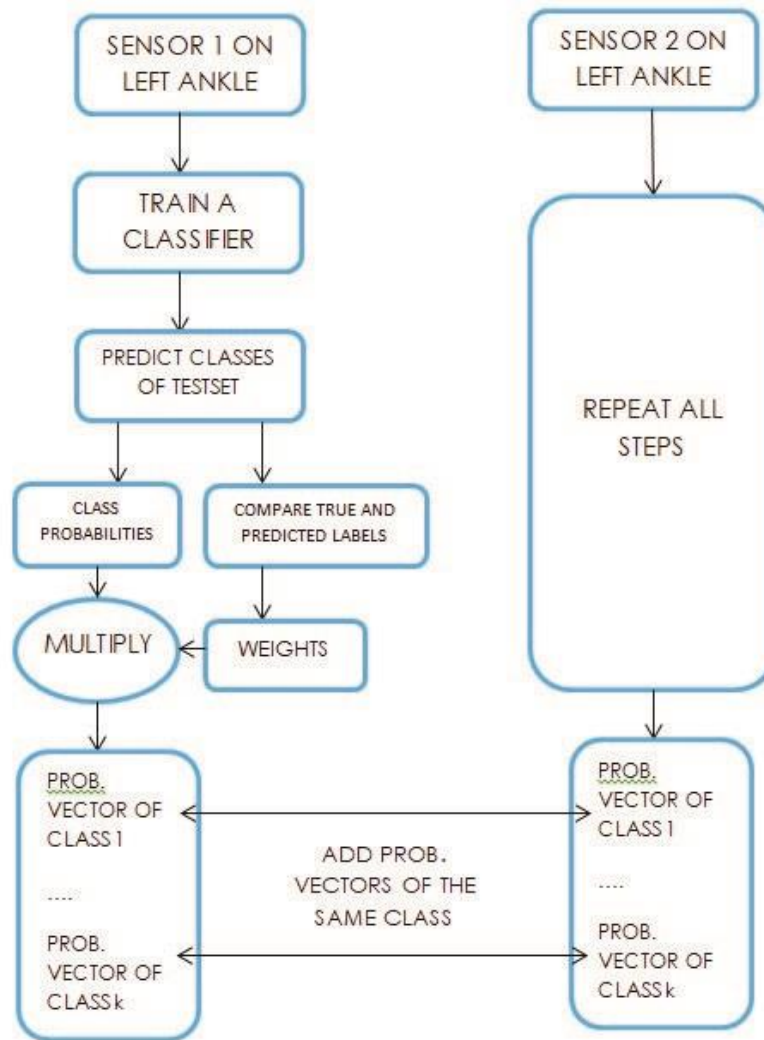


Fig.1. Example flowchart of the weighted fusion of two sensors.

SVM applied on gyroscope of the left ankle, does not again recognize "cycling", while kNN on gyroscope data does not predict "standing still". Three activities were not detected from the worst performing classifier, SVM, on magnetometer of the left ankle, namely "lying down", "running" and "jogging". It is obvious that each sensor provides better recognition of some activities, therefore the fusion of all of them is expected to utilize all information in order to detect all performed activities. SVM, however, failed to detect "cycling" on both fusion methods, which may be an indication that the sensors need to be placed elsewhere to better recognize that activity.

The accelerometer of the right lower arm failed to recognize different activities than the accelerometer placed on the left ankle. More specifically, SVM, RF and C5 do not predict "walking", while SVM still does not recognize "cycling". "Lying down" is the activity that gyroscope of the right lower arm cannot recognize with three tested classifiers: SVM, RF and C5. SVM does not detect "walking" and "cycling" also. SVM on magnetometer data does not detect "lying", "jogging", "cycling" and "running". Two activities are not predicted from the model-based fusion method when using SVM, while RF, which in general performed very well in most cases, does not predict "walking" from the class-based fusion. The results reveal that right lower arm may probably be a worst spot than left ankle to place those sensors.

Following are the results of the four classifiers applied on each sensor separately for each location and the results of fusion of the same sensors of all placements. As it can be seen in Table 4, in most algorithms, the best results are obtained from the accelerometer when placed on the right lower arm. Fusion, whether model or class based, did not improve the recognition rate in all classification algorithms except SVM.

Table 4. Fusion of accelerometers of different placements

	SVM		RF		C5		KNN	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Accel, chest	0.6458	0.5721	0.7792	0.7255	0.7208	0.6379	0.6208	0.5726
Accel, left ankle	0.5750	0.5004	0.7750	0.7670	0.5667	0.5649	0.7500	0.7489
Accel, right lower arm	0.6417	0.5582	0.8750	0.8333	0.8750	0.8333	0.9000	0.8845
WACC	0.6833	0.6202	0.8042	0.7563	0.8375	0.7955	0.8708	0.8565
DR	0.6833	0.6211	0.7792	0.7210	0.7458	0.6627	0.7667	0.7498
Averaging	0.6750	0.6132	0.7792	0.7210	0.7458	0.6627	0.7667	0.7498

Table 5 shows the results for gyroscope. Gyroscope seems to perform better when placed on the left ankle. Here, fusion of gyroscopes of different placements, improves the recognition rate compared to that of individual gyroscopes.

In general, fusion of gyroscopes results in better recognition than the fusion of accelerometers, an indication that these activities are better detected by a gyroscope. However in some cases, there were still activities not detected at all, like "lying", that was not recognized by gyroscopes and three algorithms, kNN, RF and SVM when using class-based fusion.

Table 5. Fusion of gyroscopes of different placements

	SVM		RF		C5		KNN	
	ACC	F1	ACC	F1	ACC	F1	ACC	F1
Gyro, left ankle	0.7000	0.6900	1.0000	1.0000	0.9542	0.9537	0.8583	0.8167
Gyro, right lower arm	0.5292	0.4200	0.8500	0.8085	0.8375	0.7978	0.7083	0.6823
WACC	0.6417	0.5626	0.8750	0.8333	0.9958	0.9958	0.8708	0.8292
DR	0.6375	0.5621	0.8750	0.8333	0.8708	0.8270	0.8042	0.7669
Averaging	0.6458	0.5606	0.8750	0.8333	0.9958	0.9958	0.8667	0.8249

5 Conclusion

The fusion of the three sensors improved the recognition rate, whether the sensors were placed on the left ankle or the right lower arm. For half of the classification algorithms, class-based fusion outperformed the others. In almost all cases, model-based fusion and class-based fusion outperform the baseline method of averaging. Furthermore, for the current application, the left ankle placement achieves higher recognition rates than the right lower arm.

Fusion of the same sensors placed on different body placements did not prove so promising for the prediction of the specific activities. The fusion of accelerometers of three placements, did not exceed the individual sensor's performance in most cases, while for gyroscopes, model-based fusion with the weighted accuracy, improved the recognition rate for half of the algorithms applied. Fusion of magnetometers of different placements was not attended due to the poor performance of the sensor.

Overall, for the particular implementation, combining different sensors of the same location proved better than combining same sensor placed on different locations. The class-based fusion scheme suggested in [16] performed equally well with the model-based fusion with the use of weighted accuracy. Both fusion schemes outperform fusion with averaging of class probabilities. The fact that for some tests there were classes not predicted, may have affected the performance of the class-based fusion.

For future work, the authors will investigate fusion frameworks that combine different sensors and different placements, that will eliminate the heterogeneity caused by both factors.

References

1. Chernbumroong, Saisakul and Cang, Shuang and Yu, Hongnian: Genetic algorithmbased classifiers fusion for multisensor activity recognition of elderly people. *IEEE journal of biomedical and health informatics* **19**(1), 282–289 (2014)
2. Chen, Liming and Hoey, Jesse and Nugent, Chris D and Cook, Diane J and Yu,Zhiwen: Sensor-based activity recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **42**(6), 790–808 (2012)
3. Luštrek, Mitja and Kaluža, Boštjan: Fall detection and activity recognition withmachine learning. *Informatica* **33**(2), (2009)
4. Shoaib, Muhammad and Bosch, Stephan and Incel, Ozlem and Scholten, Hans andHavinga, Paul: Fusion of smartphone motion sensors for physical activity recognition. *Sensors* **14**(6), 10146–10176 (2014)
5. Chowdhury, Alok Kumar and Tjondronegoro, Dian and Chandran, Vinod and Trost,Stewart G: Physical activity recognition using posterior-adapted class-based fusion of multiaccelerometer data. *IEEE journal of biomedical and health informatics* **22**(3), 678–685 (2017)
6. Bao, L., & Intille, S. S.: Activity recognition from user-annotated acceleration data.In: Springer International conference on pervasive computing, pp. 1–17. Springer, Berlin, Heidelberg (2004).
7. Cleland, Ian, et al.: Optimal placement of accelerometers for the detection of everyday activities. *Sensors* **13**(7), 9183-9200 (2013)
8. Zhu, C., Sheng, W.: Human daily activity recognition in robot-assisted living using multi-sensor fusion. In: 2009 IEEE International Conference on Robotics and Automation, pp. 2154-2159. IEEE, (2009)
9. Tabar, A. M., Keshavarz, A., Aghajan, H.: Smart home care network using sensorfusion and distributed vision-based reasoning. In: Proceedings of the 4th ACM international workshop on Video surveillance and sensor networks, pp. 145-154. ACM, (2006)
10. Crispim-Junior, C. F., Ma, Q., Fosty, B., Romdhane, R., Bremond, F., & Thonnat, M.: Combining multiple sensors for event detection of older people. *Health Monitoring and Personalized Feedback using Multimedia Data*, 179–194 (2015)
11. Piquier, J., Karaman, S., Letoupin, L., Guyot, P., Mgret, R., Benois-Pineau, J., ... Dartigues, J. F.: Strategies for multiple feature fusion with hierarchical hmm: application to activity recognition from wearable audiovisual sensors. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), pp. 3192-3195. IEEE, (2012)
12. Guiry, J., Van de Ven, P., & Nelson, J.: Multi-sensor fusion for enhanced contextual awareness of everyday activities with ubiquitous devices. *Sensors* **14**(3), 5687-5701 (2014)
13. Stein, S., McKenna, S. J.: Combining embedded accelerometers with computer vision for recognizing food preparation activities. In: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp. 729-738. ACM, (2013)
14. Liparas, D., HaCohen-Kerner, Y., Moutzidou, A., Vrochidis, S., & Kompatsiaris,I.: News articles classification using Random Forests and weighted multimodal features. In: Information Retrieval Facility Conference, pp. 63-75. Springer, (2014)
15. Banos, O., Garcia, R., Holgado-Terriza, J. A., Damas, M., Pomares, H., Rojas,I., ... & Villalonga, C.: mHealthDroid: a novel framework for agile development of mobile health applications. In: International workshop on ambient assisted living, pp. 91-98. Springer, (2014)
16. Author , Author: Paper title. In: Proceedings. Editor, (Year)
17. Jain, A., & Kanhangad, V. : Human activity classification in smartphones using accelerometer and gyroscope sensors. *IEEE Sensors Journal* **18**(3), 1169-1177 (2017)
18. Banos, O., Galvez, J. M., Damas, M., Pomares, H., Rojas, I.: Window size impact in human activity recognition. *Sensors* **14**(4), 6474-6499 (2014)