A Weighted Late Fusion Framework for Recognizing Human Activity from Wearable Sensors

Athina Tsanousa
Information Technologies Institute
Centre for Research and Technology
Hellas, Thessaloniki, Greece
atsan@iti.gr

Georgios Meditskos
Information Technologies Institute
Centre for Research and Technology
Hellas, Thessaloniki, Greece
gmeditsk@iti.gr

Stefanos Vrochidis
Information Technologies Institute
Centre for Research and Technology
Hellas, Thessaloniki, Greece
stefanos@iti.gr

Ioannis Kompatsiaris
Information Technologies Institute
Centre for Research and Technology
Hellas, Thessaloniki, Greece
ikom@iti.gr

***Abstract*—Following the technological advancement and the constantly emerging assisted living applications, sensor-based activity recognition research receives great attention. Until recently, the majority of relevant research involved extracting knowledge out of single modalities, however, when individual sensors performances are not satisfactory, combining information from multiple sensors can be of use and improve the activity recognition rate. Early and late fusion classifier strategies are usually employed to successfully merge multiple sensors. This paper proposes a novel framework for combining accelerometers and gyroscopes at decision level, in order to recognize human activity. More specifically, we propose a weighted late fusion framework that utilizes the detection rate of a classifier. Furthermore, we propose the modification of an already existing class-based weighted late fusion framework. Experimental results on a publicly available and widely used dataset demonstrated that the combination of accelerometer and gyroscope under the proposed frameworks improves the classification performance.**

***Index Terms*—human activity recognition, late fusion, multimodal data, accelerometers, gyroscopes**

I. Introduction

Activity recognition is a quite critical task, involved in many applications in health, technology and even security. Such applications are not only focused on the recognition of activity in the exact sense, such as walking and cooking, but might also concern harmful event detection, like fall [1], motion gestures recognition [2] and even emotion recognition [3]. Human activity recognition problems in particular can be categorized as vision-based, sensor-based [4] or combine both categories of input. In summary, a human activity recognition framework begins with the collection or extraction of raw data, like sensor signals or video images. Especially in sensor-based activity recognition research, initial data need to be preprocessed so as to eliminate diversity. Preprocessing methods include filtering and normalization. An appropriate time window is afterwards selected in order to extract features from the preprocessed data. Feature selection techniques can be utilized to select the most suitable extracted features that will then enter a classifier model so as to recognize the activities conducted [5].

Accelerometers are the sensors most broadly used since they have proven to be very effective in recognizing activities, especially the ones with repetitive body motion [23]. Accelerometers capture the magnitude and direction of an object in motion. However when used alone, they lack the ability to recognize similar activities [15], therefore combining them with other wearable or ambient sensors, improves the performance of a system. Gyroscopes measure the angular velocity of rotation [15] hence they detect the object's orientation [24]. They are also found in most activity recognition studies, however they are not so often used individually. The combination of these sensors at any level, will most probably improve the performance of a recognition system, since one sensor may cover for the deficiencies of the other e.g. the rotation speed provided by gyroscopes can correct accelerometer errors [17]. Both sensors are embedded in all smartphones and smartwatches, which makes it easier to obtain their sensor readings. Most devices have triaxial sensors, that produce three component vectors of raw signals, with each component responding to an axis of the Cartesian reference system [22].

Numerous machine learning algorithms are employed in activity recognition problems, which usually involve multilabel classification. The choice is subjective and the classifier's performance is affected by many factors like the nature of the activities performed, the type of data and the selected features. However, there are algorithms found to achieve high performance rates across many studies, like Support Vector Machines (SVM), Naive Bayes (NB) and Decision Trees [6].

As already mentioned, the existence of multiple sensors can lead to a better recognition rate when individual sensors' performances are not satisfying. Combination of multiple sensors can be achieved through fusion, early or late. Early fusion refers to combination of features, while late fusion refers to the combination of results. The most common early fusion technique is concatenation of feature vectors. Some basic late fusion strategies are a) averaging the predicted class probabilities of multiple models and b) the majority voting, which assigns to a case the label predicted by most of the models used [7]. Variants of the aforementioned strategies are the weighted averaging and weighted voting, with weights assigned to the classifiers results according to a criterion, which is usually the performance of the classifier [8]. Some of the more complex late fusion techniques are bagging, boosting and stacking [16]. Weights can be incorporated in numerous fusion strategies usually enhancing the results of the classifiers that perform best.

In the current study, we propose a weighted late fusion framework, with weights based on the detection rate (DR) of each activity. The class prediction probabilities of each sensor are weighted with the supplementary of the corresponding detection rate and then the weighted results of the two sensors are combined. From empirical experimentation, the proposed framework was found to improve the recognition rates of the late fusion implementation. Detection rate reflects the ratio only of the true positive (TP) findings of a classifier, thus its value is usually smaller than other more widely used evaluation metrics. To the best of our knowledge, the detection rate has not been utilized in fusion applications yet, at least in the activity recognition field. Furthermore, we suggest modifying the class-based weighted late fusion framework proposed in [9]. The authors in [9] used class-based weights that reflect the performance of the classifier on the training set, in terms of the F1-score. These weights are then adjusted with the class probabilities obtained from the prediction on the test set. We suggest replacing the F1-score in the calculation of the class-based weights with the detection rate. The novelty of this work can be summarized in the following:

1) the suggestion of a novel weighted late fusion framework for the combination of accelerometers and gyroscopes
2) utilizing detection rates in weighted fusion
3) the modification of an existing class-based weighted framework for late fusion

The proposed fusion schemes were evaluated on a publicly available dataset, the HAR dataset [18]. The HAR dataset consists of wearable sensors' data that recorded 6 daily activities of 30 subjects. The sensors were embedded on a smartphone mounted on the subjects' waist. The activities performed were: walking, walking upstairs, walking downstairs, sitting, standing and laying down. Four classifiers, widely applied on multilabel data for human activity recognition, were used: Random Forests, C5, kNN and Adaboost.

As far as recognition datasets is concerned, HAR is one of the most known and widely used. It was introduced in [18], where the authors tested only one algorithm, the multiclass SVM, using all extracted features of the available sensors, resulting in 96% accuracy. Since then, HAR dataset has been utilized in numerous works, many of which apply deep learning, that generally results in higher accuracy rates than machine learning algorithms, however it is time and source consuming. In [19], using a subset of the features included in HAR dataset, five ensemble classifiers were used to combine results of the two base learners, SVM and Random Forests. To overcome issues of overfitting, [20] propose a deep convolutional neural network, the perceptionNet, for late fusion and utilized the HAR [18] dataset to tune the hyperparameters.

The rest of the paper is organized as follows: in Section 2 an overview of related work is presented. In Section 3, the proposed frameworks are described, while Section 4 includes the description of the experimental setup, followed by the application and experimental results of the suggested frameworks. Finally, in Section 5 the conclusions of this work are briefly discussed.

II. Related Work

Early or late fusion is used in activity recognition to fuse features or results from different sensors, or even sensors placed on different locations. A thorough overview of fusion methods for human activity recognition from wearable sensors can be found in [14]. The authors describe several techniques for data, feature and late fusion and discuss the strengths and weaknesses of different combinations of sensors. Fusion of accelerometer and gyroscope data is the combination most commonly found in relevant studies. In [10], the authors propose the use of two descriptors in order to extract feature sets from accelerometer and gyroscope signals. They compare the results of feature and late fusion, resulting in better performance of feature fusion, which is conducted by simple concatenation of the extracted feature sets. A data fusion approach is presented in [26], that combines data from accelerometers and gyroscopes to classify daily activities and predict falls. The proposed classification algorithm uses a threshold mechanism to combine features from the two sensors. Convolutional neural networks are employed in [27] to fuse accelerometer and gyroscope data at different stages of the network. Different stages within the deep learning algorithm respond to different types of fusion. The authors concluded that in their application, late and hybrid fusion perform better than early one. In the current application, accelerometers and gyroscopes are combined using two late fusion frameworks, a proposed weighted late fusion with weights based on detection rate and a modified weighted late fusion framework.

Concatenation of feature vectors is probably the most frequent practice of fusion at feature level. Concatenation is even found in early fusion of quite heterogeneous sensors. In [25] feature vectors from a variety of sensors, like wearable ones (accelerometers and magnetometers), location and temperature sensors, were simply put together and an

”one-vs-one” approach was followed to recognize activities. To eliminate variability due to the diverse nature of the variables, authors normalized the data before training. In [11] concatenation is employed to create various sets of features derived from three sensors, namely accelerometers, gyroscopes and magnetometers, and later use these concatenated features in three types of artificial neural networks (ANN). In this work, we chose to combine the accelerometer and gyroscope sensors on a decision level instead of just concatenating features of different nature.

Late fusion allows for more experimentation and development of novel algorithms beyond the state-of-the-art. In [9] the authors combine the results of accelerometers placed at different body locations with model-based and class-based weighted decision fusion techniques and also propose a posterior adaptation of the class-based scheme. Our work suggests a modification of that proposed framework, by a different calculation of the class-based weights. In [12] they apply two fusion techniques, hierarchical decision and majority voting and introduce a novel one, the hierarchical-weighted classification. The proposed method combines the benefits of the aforementioned established fusion techniques and by using weights reflecting each entity’s performance, they create a ranking system for the importance of each component to the final hierarchical fusion scheme. Reference [21] applies several late fusion and weighted late fusion methods on multimodal data to classify 13 activities and concluded that among sensors, accelerometer and gyroscope were the most important for classifying the activities. Six different weights are incorporated in late fusion. The definition of weights reflects the performance of the models. Accuracy and mean square error are used for their calculation. Our proposed frameworks introduce the use of detection rate in order to evaluate the performance of models.

III. Methodology

In this section we describe the proposed frameworks for the recognition of activities from multisensor data. Firstly we propose a novel way to combine results of many sensors by a weighted late fusion framework that uses weights related to the detection rate of each class. Secondly, we introduce a variation of the class-based weighted late fusion framework proposed in [9].

A. Weighted late fusion framework

Consider a multilabel classification problem of *k* classes (i.e. activities) and *m* models, where each model corresponds to a different sensor. The goal is to combine the results of these models in such a way that the recognition of the classes is improved.

Classification problems consist of a training and a testing stage. During the training stage of a model, the classifier is trained on the features of each sensor using a 10-fold cross validation. Proceeding with the testing stage, the trained model outputs for each test case a) a predicted label and b) a probability score *P(x)*, expressing how possible it is for each test case to belong to a class. In order to utilize the information provided by the *m* models, we suggest combining the probability vectors (1) of different models with weighted late fusion.

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

Each model will be assigned weights (2) that relate to the classifier’s ability to detect true positive (TP) cases among all predictions, which is expressed by the detection rate of a class. Detection rate, defined in (3), is considered a strict evaluation metric since it focuses on the discovery of the true positives and not all true findings of an algorithm. All evaluation metrics are generally obtained when the labels predicted by the classifier are compared with the true classes. For multiclass problems, the comparison of predicted and actual classes is done with the one vs all approach, meaning that the class to be evaluated consists the ”positive” findings and all the rest the ”negative” findings [13]. In some papers, the term detection rate refers to the recall/sensitivity [28], which still measures the detection of the true positives but among all the positive cases only, i.e. true positives and false negatives (TP+FN), while the current detection rate refers to the ratio of true positives among all findings, including true negatives (TN) and false positives (FP) too.

$$W_{ij} = \{w_{i1}, ..., w_{ik}\} \tag{2}$$

$$DR = TP/(TP + TN + FP + FN) \tag{3}$$

To assist the recognition of classes not so easily detected, we set the weights equal to the supplementary of the detection rate (4).

$$W_{ij} = 1 - DR_{ij} \tag{4}$$

Weights are calculated for each class and are then multiplied by the corresponding probability vectors as in (5). For each class there will be *m* weighted probability vectors, each one corresponding to a different sensor. The weighted probabilities P_w of the m models will be summed together using (6), forming a final score for each class. The final predicted label of each test case is the class with the maximum final score. The proposed framework is graphically presented in Fig. 1.

$$P_w = W_{ij} P_{ij} \tag{5}$$

$$Score_j(x) = \sum_i W_{ij} P_{ij} \tag{6}$$

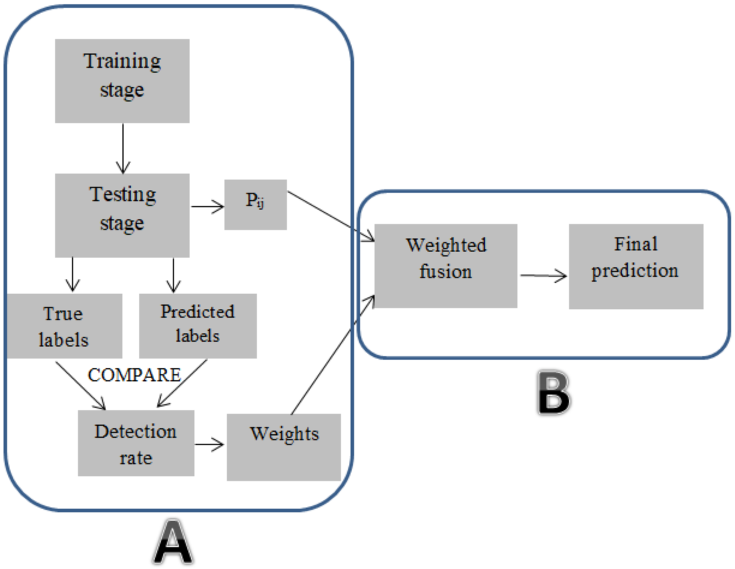

Fig. 1. Flowchart of the proposed weighted late fusion framework. Procedure in A is repeated for each sensor, while B refers to the combination of sensors.

B. Class-based weighted late fusion framework

This framework utilizes class-based weights that are based on prior knowledge of the performance of the classifier [9]. During the training stage, a model is trained using 10-fold cross-validation and then tested on the same data, i.e. the training set. The predicted labels of the train cases are compared with the true classes to evaluate the performance of the model on the train data. We suggest utilizing the detection rate for evaluating the model's performance, instead of the F1-score that is used in [9]. The formula for the calculation of detection rate is defined in (3), however the values of the metric will differ between the two frameworks, since they are obtained from different stages of the classification process.

Detection rate is then incorporated in the calculations of weights using (4). In the testing stage, the predicted class probabilities are produced. For the whole testset there will be probability vectors P_{ij} (*i=1,..,m* and *j=i,..,k*) for each class. Using an adjustment parameter *a*, the weights derived from the train set and the class probabilities of the prediction of the test set are fused using (7). The adaptation parameter *a* is assigned values ranging from 0 to 1 [9].

$$AP_{ij}(x) = aW_{ij} + (1-a)P_{ij} \tag{7}$$

Proceeding with the fusion, the final weighted class probabilities of each model are added together, using (8) to form a vector of scores for each class. This results in each test case having a vector of scores corresponding to each class. The class with the maximum score is assigned as the final predicted label for each test case. The modified framework is illustrated in Fig. 2.

$$Score_j(x) = \sum_i AP_{ij}(x) \tag{8}$$

IV. Evaluation

A. Experimental setup

For the evaluation of the suggested fusion scheme, the HAR dataset was chosen [18]. HAR is publicly available from the UCI Machine Learning Repository [29] and has been frequently used in the literature due to its variety of sensor signals and extracted features. The dataset consists of 30 subjects, aged 19 to 48, each one performing six activities (Walking, Walking Upstairs, Walking Downstairs, Sitting, Standing, Laying). The subjects wore a smartphone (Samsung Galaxy S II) on their waist with embedded accelerometer and gyroscope sampling at 50Hz. 70% of the obtained data were randomly chosen for training and the rest for testing the classifiers. Raw observations were filtered and a 2.56 sec sliding time window with 50% overlap was used to extract features. For more detailed information we refer the reader to the original paper [18]. Features extracted only from accelerometer and gyroscope raw data were selected to form the train and test sets of the corresponding modalities. The features used in the present analysis, as described in [18], can be found in Table I. No feature selection algorithms were applied.

The classification problem consists of the *k=6* activities and of *m=2* models, namely the accelerometer model and the gyroscope model. For the recognition of the six activities, several classifiers were tested, with the results of the four that performed better reported here, namely Random Forests, C5, kNN and Adaboost. For kNN algorithm, *k* was set to 5,7 and 9 neighbors for each round and the value that produced the optimal results was reported at the end. Each algorithm was trained using 10-fold cross-validation. In order to assess the performance of the classifiers and compare the obtained results, the overall accuracy of each algorithm is reported. Accuracy is the ratio of the correct predictions (true positives (TP) and true negatives (TN)) towards all predictions (9). For the implementation, the R package *caret* was utilized [13].

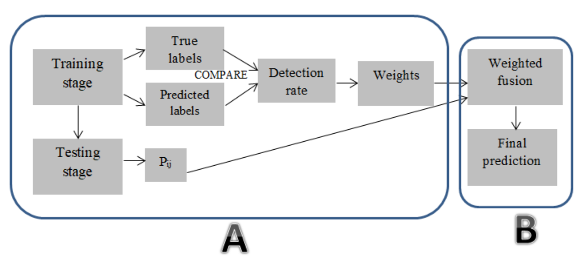

Fig. 2. Flowchart of the modified class-based weighted late fusion framework. Procedure in A is repeated for each sensor, while B refers to the combination of sensors.

TABLE I
FEATURES USED IN THE ANALYSIS

Features
Mean
Standard deviation
Median
Maximum value of the array
Minimum value of the array
Signal magnitude area
Energy
Interquartile range
Entropy
Autoregression coefficients
Correlation coefficient
Largest frequency component
Frequency signal weighted average
Skewness
Kurtosis
Energy of a frequency interval
Angle

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \quad (9)$$

To assess the performance of the proposed weighted late fusion framework, the following comparisons were made: a) with the performance of the individual sensors and b) with the performance of other well known fusion methods, i.e. averaging and stacking. The modified framework of class-based weighted late fusion was compared with the initial framework explained in [9]. The initial framework was chosen as it is quite similar to our experimental setup, combining data from accelerometers placed at different locations, while we try to combine accelerometers and gyroscopes.

B. Tests

1) Implementation of the weighted late fusion framework: In this section, we describe the application of the proposed **weighted fusion framework**. For each sensor, the trained algorithm produces the prediction probabilities of the test cases. Let *m=1* denote the model built on the accelerometer features and *m=2* the model of the gyroscope features. The probability sets of the accelerometer model (10) and the gyroscope model (11), consist of probability vectors P_{ij} (*i=1,..,m* and *j=i,..,k*), where each one contains the test cases' probabilities to be assigned to class *j*.

$$P_1 = \{P_{11}, P_{12}, ..., P_{16}\} \quad (10)$$

$$P_2 = \{P_{21}, P_{22}, ..., P_{26}\} \quad (11)$$

The detection rates of each class are obtained when comparing the predicted labels of the test cases with the actual classes. The respective values of the accelerometer model (12) and the gyroscope model (13) will be used to calculate the weights using (4). The detection rates of the four classifiers applied, were averaged for each sensor and are displayed in Table II. The activities with the maximum average detection rate are *laying* when using only the accelerometer features and *standing* when predicting with the gyroscope features.

TABLE II
AVERAGE DETECTION RATES

	Activities					
	WALK	***WU***	***WD***	***SIT***	***STAND***	***LAY***
Accel	0.1612	0.1336	0.1219	0.1256	0.1458	**0.1818**
Gyro	0.1416	0.1388	0.1066	0.1191	**0.1501**	0.1210

[a]WU stands for walking upstairs and WD for walking downstairs

$$DR_1 = \{DR_{11}, DR_{12}, ..., DR_{16}\} \quad (12)$$

$$DR_2 = \{DR_{21}, DR_{22}, ..., DR_{26}\} \quad (13)$$

Before combining the results of the two sensors, the probability vectors of each model need to be multiplied by the corresponding weights using (5), resulting in two vectors of weighted probabilities for each sensor. The weighted probabilities of the two sensors are finally added together to form a final score for each class. Classes with the maximum score are assigned as final labels to each test case. The described procedure is graphically depicted in Fig. 3.

The comparison of individual sensor's performance and the proposed method (Table III) revealed superiority of the proposed fusion framework for all four classification algorithms tested. Table IV shows the comparison of predicted and true labels of the proposed weighted fusion framework, with classifier C5, that performed better among the four algorithms. *Walking, Walking upstairs* and *Laying* were the activities better recognized.

Regarding the recognition rate of individual activities over all four classifiers in Table V, *Laying* was the activity with the highest rate, while *sitting* has the smallest value over the four classifiers.

The results of the proposed weighted late fusion framework were also compared with the results of other popular late fusion techniques. Particularly, we applied averaging of the class probabilities and stacking with two algorithms: a) SVM, which is widely used as a base learner in activity recognition problems and b) Gradient Boosting Machine (GBM), a boosting algorithm that is usually employed in stacking technique. For averaging, the class probabilities of accelerometer and gyroscope models are averaged and the class with the

TABLE III
COMPARISON OF RESULTS OF INDIVIDUAL SENSORS AND PROPOSED FRAMEWORK

	Accelerometer	***Gyroscope***	***Weighted late fusion***
Random Forests	0.8697	0.8208	**0.9277**
C5	0.8833	0.8161	**0.9294**
kNN	0.8588	0.7241	**0.8972**
Adaboost	0.8677	0.7479	**0.8996**

[a]The cells include the accuracy values

TABLE IV
CONFUSION MATRIX OF C5

	Activities					
	WALK	*WU*	*WD*	*SIT*	*STAND*	*LAY*
WALK	487	4	9	0	0	0
WU	5	461	33	0	0	0
WD	4	6	378	0	0	0
SIT	0	0	0	413	66	3
STAND	0	0	0	78	466	0
LAY	0	0	0	0	0	534

TABLE V
AVERAGE BALANCED ACCURACY OVER FOUR CLASSIFIERS

	Activities					
	WALK	*WU*	*WD*	*SIT*	*STAND*	*LAY*
Average Accuracy	0.9793	0.958	0.9346	0.897	0.9204	0.9945

highest averaged probability is assigned to every test case. Stacking trains the selected algorithm on the predicted class probabilities of other base learners. Here, the base learners are the accelerometer and gyroscope models. As shown in Table VI, the proposed framework outperforms most of the other fusion techniques.

2) Implementation of the class-based weighted late fusion framework: Following is the application of the modified class-based fusion framework on the HAR dataset. During the training stage of each classification algorithm, the trained models were used to output predictions on the same data they were trained on. The performance of the selected algorithms was evaluated using the detection rate and weights were calculated again using formula (4). In the testing stage, the trained model was used to predict the labels and produced the class probabilities P_{ij} (*i=1,2* and *j=1,..,6*). The posterior probabilities obtained from the prediction on the testset were combined with the class-based weights using (7). The value of adaptation parameter was set to 0.25 since it produced the optimal results. The described procedure was repeated separately for accelerometer and gyroscope features (Fig. 4). The comparison of the original framework and the proposed modification (Table VII) shows that the modified framework outperforms the original in three of the four classification algorithms used.

TABLE VI
COMPARISON OF THE PROPOSED FRAMEWORK AND OTHER FUSION METHODS

	Weighted Late Fusion	*Averaging*	*SVM Stacking*	*GBM Stacking*
Random Forests	**0.9277**	0.9267	0.7978	0.9165
C5	0.9294	**0.9298**	0.8235	0.9158
kNN	0.8972	0.8918	0.7869	**0.9036**
Adaboost	**0.8996**	0.8966	0.6047	0.8278

[a]The cells include the accuracy values

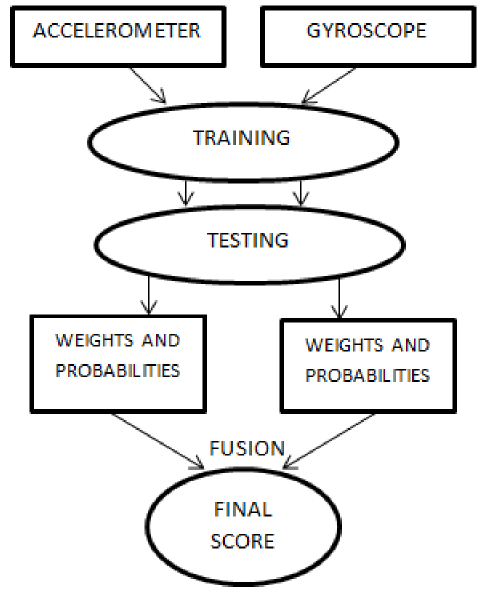

Fig. 3. Implementation of weighted late fusion

TABLE VII
COMPARISON OF ORIGINAL AND MODIFIED CLASS-BASED WEIGHTED LATE FUSION

Classifier	*Original framework*	*Modified framework*
Random Forests	0.9186	**0.927**
C5	0.7479	**0.9304**
KNN	**0.8979**	0.8958
Adaboost	0.8992	**0.9006**

[a]The cells include the accuracy values

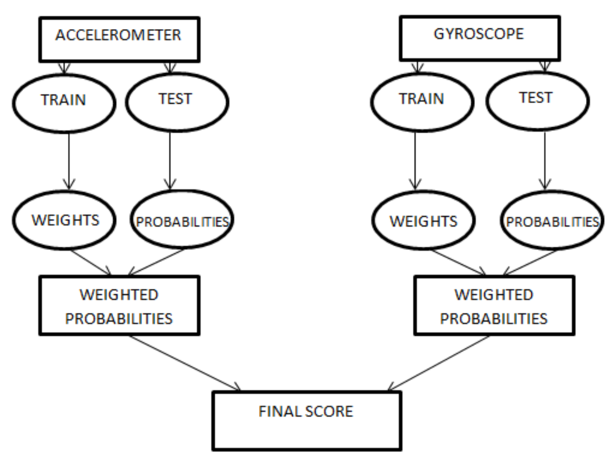

Fig. 4. Implementation of the modified class-based weighted late fusion

V. Conclusions

The combination of multiple sensors assists in improving the recognition of multiple activities. Although accelerometers and gyroscopes are usually combined on feature level with simple concatenation, here we suggested decision fusion of those sensors for a multiclass activity recognition problem. We proposed a weighted late fusion strategy for combining the classification results of individual sensors and we incorporated the detection rate of a classifier for the calculation of weights. Detection rate is a performance evaluation metric that hasn't been employed to weighted frameworks to the extent of our knowledge. Furthermore, using weights based on the class detection rate, we suggested a variation of a class-based weighted fusion strategy.

Four classifiers were used to evaluate the proposed frameworks, with C5 and Random Forests achieving the higher recognition rates. The experimental results revealed superiority of the proposed scheme for the majority of the comparisons conducted for both frameworks. To the extent of our knowledge, detection rate has not been utilized in weighted fusion schemes, especially in the activity recognition literature, and it could constitute an alternative solution for late fusion.

Suggestions for future work include utilizing the proposed frameworks in other application fields as well as incorporating detection rate in more complex weighting schemes. An indicative application could be to combine heterogeneous sensors for human localization. Other suggestions include detection of harmful events, since different data sources are utilized and fusion is a suitable method for the exploitation of all information available. The proposed frameworks will be tested in the future in a real world clinical environment and a smart home.

Acknowledgment

This research has been cofinanced by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship and Innovation, under the call RESEARCH - CREATE - INNOVATE (project code:T1EDK-00686) and by beAWARE project partially funded by the European Commission under grant agreement No 700475.

References

[1] Zhang, T., Wang, J., Xu, L., and Liu, P. (2006). Fall detection by wearable sensor and one-class SVM algorithm. In Intelligent computing in signal processing and pattern recognition (pp. 858-863). Springer, Berlin, Heidelberg.

[2] Kratz, S., Rohs, M., and Essl, G. (2013, March). Combining acceleration and gyroscope data for motion gesture recognition using classifiers with dimensionality constraints. In Proceedings of the 2013 international conference on Intelligent user interfaces (pp. 173-178). ACM.

[3] Poria, S., Chaturvedi, I., Cambria, E., and Hussain, A. (2016, December). Convolutional MKL based multimodal emotion recognition and sentiment analysis. In 2016 IEEE 16th international conference on data mining (ICDM) (pp. 439-448). IEEE.

[4] Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., and Yu, Z. (2012). Sensor-based activity recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6), 790-808.

[5] Stisen, A., Blunck, H., Bhattacharya, S., Prentow, T. S., Kjrgaard, M. B., Dey, A., ... and Jensen, M. M. (2015, November). Smart devices are different: Assessing and mitigating mobile sensing heterogeneities for activity recognition. In Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems (pp. 127-140). ACM.

[6] Lara, O. D., and Labrador, M. A. (2013). A survey on human activity recognition using wearable sensors. IEEE Communications Surveys and Tutorials, 15(3), 1192-1209.

[7] Kuncheva, L. I. (2002). A theoretical study on six classifier fusion strategies. IEEE Transactions on pattern analysis and machine intelligence, 24(2), 281-286.

[8] Mangai, U. G., Samanta, S., Das, S., and Chowdhury, P. R. (2010). A survey of decision fusion and feature fusion strategies for pattern classification. IETE Technical review, 27(4), 293-307.

[9] Chowdhury, A. K., Tjondronegoro, D., Chandran, V., and Trost, S. G. (2018). Physical Activity Recognition Using Posterior-Adapted Class-Based Fusion of Multiaccelerometer Data. IEEE journal of biomedical and health informatics, 22(3), 678-685.

[10] Jain, A., and Kanhangad, V. (2018). Human Activity Classification in Smartphones Using Accelerometer and Gyroscope Sensors. IEEE Sensors Journal, 18(3), 1169-1177.

[11] Pires, I. M., Garcia, N. M., Pombo, N., Flrez-Revuelta, F., Spinsante, S., and Teixeira, M. C. (2018). Identification of activities of daily living through data fusion on motion and magnetic sensors embedded on mobile devices. Pervasive and Mobile Computing, 47, 78-93.

[12] Banos, O., Damas, M., Pomares, H., Rojas, F., Delgado-Marquez, B., and Valenzuela, O. (2013). Human activity recognition based on a sensor weighting hierarchical classifier. Soft Computing, 17(2), 333-343.

[13] Kuhn, M. (2008). Building predictive models in R using the caret package. Journal of statistical software, 28(5), 1-26.

[14] Nweke, H. F., Teh, Y. W., Mujtaba, G., and Al-Garadi, M. A. (2019). Data fusion and multiple classifier systems for human activity detection and health monitoring: Review and open research directions. Information Fusion, 46, 147-170.

[15] Wang, A., Chen, G., Wu, X., Liu, L., An, N., and Chang, C. Y. (2018). Towards Human Activity Recognition: A Hierarchical Feature Selection Framework. Sensors, 18(11), 3629.

[16] Su, X., Tong, H., and Ji, P. (2014). Activity recognition with smartphone sensors. Tsinghua science and technology, 19(3), 235-249.

[17] Ustev, Y. E., Durmaz Incel, O., and Ersoy, C. (2013, September). User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal. In Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication (pp. 1427-1436). ACM.

[18] Anguita, D., Ghio, A., Oneto, L., Parra, X., and Reyes-Ortiz, J. L. (2013, April). A public domain dataset for human activity recognition using smartphones. In ESANN.

[19] Elamvazuthi, I., Izhar, L., and Capi, G. (2018). Classification of Human Daily Activities Using Ensemble Methods Based on Smartphone Inertial Sensors. Sensors, 18(12), 4132.

[20] Kasnesis, P., Patrikakis, C. Z., and Venieris, I. S. (2018, September). PerceptionNet: A Deep Convolutional Neural Network for Late Sensor Fusion. In Proceedings of SAI Intelligent Systems Conference (pp. 101-119). Springer, Cham.

[21] Chernbumroong, S., Cang, S., and Yu, H. (2015). Genetic algorithm-based classifiers fusion for multisensor activity recognition of elderly people. IEEE journal of biomedical and health informatics, 19(1), 282-289.

[22] Brezmes, T., Gorricho, J. L., and Cotrina, J. (2009, June). Activity recognition from accelerometer data on a mobile phone. In International Work-Conference on Artificial Neural Networks (pp. 796-799). Springer, Berlin, Heidelberg.

[23] Chen, L., Hoey, J., Nugent, C. D., Cook, D. J., and Yu, Z. (2012). Sensor-based activity recognition. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 42(6), 790-808.

[24] Lutrek, M., and Kalua, B. (2009). Fall detection and activity recognition with machine learning. Informatica, 33(2).

[25] Fleury, A., Vacher, M., and Noury, N. (2010). SVM-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results. IEEE transactions on information technology in biomedicine, 14(2), 274-283.

[26] Ando, B., Baglio, S., Lombardo, C. O., and Marletta, V. (2016). A multisensor data-fusion approach for ADL and fall classification. IEEE Transactions on Instrumentation and Measurement, 65(9), 1960-1967.

[27] Muenzner, S., Schmidt, P., Reiss, A., Hanselmann, M., Stiefelhagen, R., and Drichen, R. (2017, September). CNN-based sensor fusion techniques for multimodal human activity recognition. In Proceedings of the 2017 ACM International Symposium on Wearable Computers (pp. 158-165). ACM.

[28] Godil, A., Bostelman, R., Shackleford, W., Hong, T., and Shneier, M. (2014). Performance metrics for evaluating object and human detection and tracking systems (No. NIST Interagency/Internal Report (NISTIR)-7972).

[29] Lichman, M. UCI Machine Learning Repository. Available online: http://archive.ics.uci.edu/ml (accessed on 19 May 2017).