2019 IEEE World Congress on Services (SERVICES)

On the Homogenization of Heterogeneous Inertial-based Databases for Human Activity Recognition

Anna Ferrari, Daniela Micucci, Marco Mobilio and Paolo Napoletano
Department of Informatics, Systems, and Communication
University of Milano - Bicocca, Milan, Italy
{*anna.ferrari,daniela.micucci,marco.mobilio,paolo.napoletano*}*@unimib.it*

***Abstract*—In the last years supervised machine learning techniques are largely employed for automatic Human Activity Recognition (HAR) using inertial sensors, such as accelerometer and gyroscope. HAR has many applications in several domains such as, for example, healthcare, sport, and entertainment. Machine learning scientists made available to the community a plenty of labeled databases for benchmarking that, unfortunately, are not consistent, both syntactically (e.g., different sampling frequency) and semantically (e.g., labels with different meanings). Commonly, due to this inconsistency, scientists evaluate their progress on individual databases separately, which corresponds to training and testing using the same database. Coherent merging of existing databases would enable: 1) evaluation of generalization capabilities of methods across databases; 2) use of deep learning techniques that, unlike traditional ones, require much more labeled data for the training process. Moreover, the growth in the daily use of wearable devices will produce a big amount of inertial data which, if not correctly labeled, cannot be efficiently exploited for the study of automatic HAR. In this paper we propose a semi-automatic procedure to coherently merge existing databases based on signal and word similarity. Preliminary experiments demonstrates the effectiveness of the proposed procedure.**

***Keywords*-Big labeled data; Deep Learning; Human Activity Recognition;**

I. Introduction

In recent years, Human Activity Recognition (HAR), has become a very active research field with applications in many different areas, such as medical, security, entertainment, and tactical scenarios [1].

HAR techniques should be able to automatically recognize human activities, including possibly falls [2], relying of samples acquired by devices. Although the use of cameras to automatically detect activities is very common [3], in recent years inertial sensors have started to be extensively used as acquisition devices because of their non-intrusiveness, cost-effectiveness, and ease of availability [1].

The application of machine learning based techniques, and especially deep learning techniques, to perform HAR is growing rapidly. In this context, the databases used to train the classifiers are of great importance. In order to avoid overfitting in large size networks and to achieve good performances, large scale databases are needed [4], [5], [6], [7]. In order to fulfil this necessity, many databases have been collected and some of them have been made publicly available, such as those in the UCI Repository [8].

Most of the available databases, however, are not suitable for deep learning because of their small dimensions. Merging of multiple databases would be a solution to produce a larger one. However, merging databases presents many challenges and open issues.

In this paper we focus on the problems related to inconsistencies that may raise due to *syntactic* and *semantic* inconsistencies. Syntactic issues are related to the variability of the device adopted for recording the data, for instance an accelerometer sensor may acquire at 50Hz and another at 200Hz. Thus, before merging, signals need to be preprocessed in order to be compatible. Semantic issues are related to the meaning of the words adopted by scientists to label the human activities. The set of different human activities and their variations that could be recorded is very large and complex: the same activity could be performed in many different ways, in different contexts, and with different goals.

Thus, merging apparently identical activities from different databases, is likely to result in an inconsistent set of data, that does not improve the training of a classifier.

Semantic issues are related to a not shared definition of Activities of Daily Living (ADL) that are the activites studied by the HAR community. We identified two kinds of inconsistencies related to semantic issues:

1) Different activities with the same labels (e.g., the label "stand up" is used to identify the action of going from a sitting position to a standing position in database 1, while it represents the static action of standing still in database 2).
2) The same activity addressed with different labels in each databases (e.g., the action of walking is referred as "walking" in database 1 and as "walk" in database 2).

The issue of having different labels for the same activity, may be a mere syntactic difference (like "walk" vs "walking" vs "wlk") or the result of similar concepts that results in the same practical action (e.g., the action of "jogging" in database 1 and "running" in database 2).

978-1-7281-3851-0/19/$31.00 ©2019 IEEE
DOI 10.1109/SERVICES.2019.00084

It becomes clear that careful merging strategies are necessary before being able to exploit an integrated database to train a classifier. While a simple syntactic integration can be done quickly be reviewing the labels and associating them to their counterparts, this is not enough.

In this paper we propose a procedure for merging labels that take into account the shape of the signals and similarity between words. Firstly, the similarities between signals is computed by using the nearest neighbor applied to the feature vectors that describe signals. If a pair of labels is considered similar from the signal point of view, then word similarity between labels is calculated through the Levenshtein distance in order to evaluate whether those labels can be merged or not. To the best of our knowledge no such analysis is present in the literature, where the only available toolbox aimed at unifying databases only performs syntactic analysis on the labels [9].

It is important to remark that HAR is applied in many domains, with a great number of activities that may be defined, from basic (e.g., sitting or walking) to very complex ones (e.g., washing dishes or preparing a sandwich). We focus our contribution on the basic activities of daily living related to healthcare applications. However concepts are generally valid and the approach may easily be extended to fulfill specific needs of different domains.

The paper is organized as follows: Section II provides an overview of the definitions related to human activities that are present in the literature; Section III introduces a semi-automatic procedure for avoiding misunderstanding in labels definitions; Section IV shows and discusses an application of the procedure; finally, Section V sketches some conclusions.

II. Activity Definitions in the State-of-the-art

Human activities that are usually monitored in healthcare applications are called Activities of Daily Living (ADLs). This term has been introduced first time by Katz et al. [10] with the aim of assessing through a graded index the degree of autonomy in everyday life activities of patients with chronic conditions [11], [12].

ADLs have been formally divided in BADLs (Basic ADLs), which are personal care basic activities, such as bathing, dressing, and eating, and IADLs (Instrumental ADLs), which are more complex activities, such as cooking, house cleaning, and phone usage [13], [14], [11].

Since its first introduction, the definition of ADLs has been refined many times [11], especially because the ADLs concept has been adopted in many other application domains such as sport, entertainment, and security. As a consequence, the activities, although represented with the same term, could correspond to different physical activities. To better understand the variety of activity definitions, we have searched on Google Scholar the definition of ADLs using the following keywords: “classification of human actions”, “classification of human activities”, “detection of human activities”, “human activity recognition”, “human activity recognition smartphone”, “human activity recognition inertial sensors”, and “human activity recognition accelerometer”. This research has provided thousands of results, most of which related to medical studies.

Table I shows some example of activities and their definitions as used by different scientists. As it can be noticed, there is no commonly accepted definition even for very common activities such as “walking” or “climbing stairs”. Hence the need to uniform the terminology in order to have databases with semantically homogeneous signals.

Activity	Mentioned as
walking	physical activity [15], [16], [17] dynamic activity [18], [19] ADL [20], [18], [21] simple activity [22], [23] basic and common activities in people’s daily living [24] common activities in people’s daily living [25] daily living human activities [26] basic movement [27] daily activity [28], [29] body state [30]
jumping	ADL [21] simple activity [22] physical activity [15] daily activity [29] common activities in people’s daily living [25]
walking upstairs	dynamic activity [18] simple activity [22] physical activity [17], [15] ADL [20], [18]

Table I
Examples of activity categorizations from scientists.

III. Database Homogenization

Here we describe the proposed semi-automatic procedure for the homogenization of inertial databases based on *signal* and *word similarity* matching.

Let us consider the case of two databases D_1 and D_2 made of accelerometer signals that are opportunely labeled with terms such as “walking”, “running”, “jumping”, and so on. Labels are chosen from the L_1 and L_2 dictionaries respectively.

Formally, $D_1 = \{s^1_1, \cdots, s^1_P\}$ and $D_2 = \{s^2_1, \cdots, s^2_R\}$ are the set of signals related to activities belonging to each database, and $L_1 = \{a^1_1, \cdots, a^1_N\}$ and $L_2 = \{a^2_1, \cdots, a^2_L\}$ are the dictionaries of labels used by D_1 and D_2 respectively.

The output of the homogenization procedure will be a merged database $D_M = \{s^M_1, \cdots, s^M_R\}$ containing all the signals from D_1 and D_2 that are labeled using the terms from the common dictionary $L_M = \{a^M_1, \cdots, a^M_P\}$.

The procedure can be easily extended to the case of multimodal databases that contain for instance accelerometer and gyroscope signals, as well as to the case of more than 2 databases.

The accelerometer signal is usually a temporal sequence of floating point values for the three axis x, y, z. Each temporal sequence of the databases may be of different length because of several causes. This incoherence needs to be solved through a *normalization* and *segmentation* phase, because supervised machine learning algorithms use sequences of the same size (and related labels) as input of the training process.

One reason of length incoherence is related to the sampling frequency of the devices adopted during database recording that may be different. In this case, it is necessary to re-sample all the signals in order to obtain signals at the same sampling rate. Literature suggests that about 50Hz is a suitable sampling rate that permits to model human activities [31].

Furthermore, even if the sampling rate is the same, it may happen to have signals of different temporal length. For instance a person could have walked for 1 minute and another for 1 minute and half. To cope with this problem, after the re-sampling process, each signal is segmented using a fixed size window. Literature suggests that a window of about 3 seconds is suitable to model human activities [32].

Once each database is represented as sets of fixed-size temporal sequences and related labels, our semi-automatic procedure is applied. The procedure includes three phases:

1) Signal-based similarity computation.
2) Word-based similarity computation.
3) Human expert evaluation.

A. Signal-based similarity computation

During this phase each sequence in the databases is converted into feature vectors using a method for feature extraction as described in [33]. These features are temporal and frequency statistics of the signal. The more discriminative are the features extracted from the sequences, the more accurate will be the homogenization procedure.

Once the features are extracted, a confusion matrix grouped by activities is built by calculating the Euclidean distance between each sequence. In particular, the confusion matrix of size $N \times L$ is obtained by calculating all distances between the signals belonging to all the possible pairs of activities $\{a_1^1, \cdots, a_N^1\} \times \{a_1^2, \cdots, a_L^2\}$.

Distance between pairs of activity $d(a_1^1, a_N^1)$ is used for calculating similarities between signals. The lower is the distance the higher is the similarity.

Confusion matrix is used to find similarities between pairs of activities. To this aim, the concept of *self similarity* is introduced. Self similarity is obtained by considering the distance between all the sequences of the same activity. It is used as threshold to filter out sequences that have distances higher than the self similarity. The ratio, is that the activity a_i is similar to the activity a_j if their distance is comparable with the distance of a_i with itself. The threshold is obtained as the average of the distances between all the sequences belonging to the same activity. Whatever is the database considered, each activity has its own threshold.

B. Word-based similarity computation

Once the similarity is calculated for all the activities, a list of possible similar pairs of activities is achieved by adopting the thresholds. These pairs are then analyzed by using a word processing module. All the pairs of activities with a high degree of term similarity are retained. The degree of similarity is calculated in terms of edit distance through the Levenshtein distance.

C. Human expert evaluation

The pairs that are not retained in previous steps are then evaluated by an expert. The expert analyzes the pairs of labels to understand if they might be merged. For instance, in the case of a similar pair made of "running" and "sitting", the expert decides to discard it because related to very different activities in terms of inertial movements.

In cases of more ambiguity, such as "running" vs "jogging", the expert, before accepting the merging proposed, visualizes the signals associated to sequences of the activities and he evaluates the visual similarity of the patterns.

IV. Experiments

In order to prove the effectiveness of the proposed procedure we report some preliminary experiments for merging 5 databases of accelerometer signals: UniMiB SHAR [32], MobiAct [34], Motion Sense [35], Real World [36], and UMAFall [37]. For a subset of these databases we have performed a classification experiment in order to show how much the merging procedure influences the performance of a machine learning classifier.

A. Merging procedure

By analyzing the labels used by the various databases to categorize activities and by observing the signals, we have defined a set of labels that will be used to classify the activities of the analyzed databases. Some of these labels have already been used by other databases (for example, "downstairs"), others have been coined by us because they are more expressive with respect to the activity to which they are associated (for example, "jumping fast"). These labels are shown in the column *Label* of Table II. The labels in italic are those coined by us, the others are those used in literature.

Table II shows the labels of the activities of each database before and after the merging procedure (from the second column). The columns related to each database show the original labels while the first column shows the result achieved by our procedure. From the table is clear that some activities with the same label but belonging to different databases have not been merged by the procedure. It means that during the first phase of the procedure the distances

Label	MobiAct	Motion Sense	UniMiB SHAR	Real World	UMAFall
downstairs	stairs down	downstairs	going downstairs	climbing down	go downstairs
lying still	-	-	-	lying	-
lying down	-	-	lying down from standing	-	lying down on a bed
jogging	-	-	-	running	-
jumping fast	-	-	jumping	-	-
jumping slow	jumping	-	-	jumping	-
running	jogging	jogging	running		jogging
sitting still	sitting on chair	sitting	-	sitting	-
sit to stand	sit to stand	-	standing up from sitting	-	getting up from chair
standing still	standing	standing	-	standing	-
standing up from lying	-	-	standing up from lying		
stand to sit	stand to sit	-	sitting down	-	-
upstairs	stairs up	upstairs	going upstairs	climbing up	go upstairs
walking	walking	walking	walking	walking	walking

Table II
ASSOCIATED LABELS AFTER THE SEMI-AUTOMATIC PROCEDURE

between those activities were higher than the threshold. This is the case of the activity "jumping" of the database UniMiB SHAR that has not be grouped to the activity "jumping" belonging to other database. In contrast, the activities "stairs down", "downstairs", "going downstairs", "climbing down", and "go downstairs", that belong to each of the considered database, have been merged.

B. Classification experiments

We have performed classification experiments by employing a k-NN classifier on a subset of databases: UniMiB SHAR [32], MobiAct [34], and Motion Sense [35]. Each database has been split in training and test by following a traditional 5-fold cross validation. We have measured the performance in terms of accuracy, that is, the number of sequences correctly predicted divided by the total number of sequences.

In order to understand how much the merging procedure affects the performance of the classifier, we have trained a classifier on each database independently and on the simply union of all the three databases. Then, we applied our homogenization procedure for all the databases by assigning the labels listed in column *Label* in Table II and we trained a classifier on the unified database.

Table III shows results achieved by the k-NN classifier trained on a database obtained after that our merging procedure has been applied to all the activities.

Table IV shows results achieved by the k-NN classifier trained on a database obtained after that our merging procedure has been applied to all the activities. In particular column *Original* shows the results obtained by training a classifier on each of the databases independently; column *Union* reports the accuracy obtained by a single classifier trained on a database achieved by simply unifying all the three databases without merging the activities; and column *Merging* shows how much the classifier performance is improved with respect to the classifier trained on the database achieved by unifying the labels.

Number	Label	Accuracy
1	walking	0.9721
2	running	0.9445
3	upstairs	0.5635
4	downstairs	0.5290
5	standing still	0.6374
6	jumping fast	0.3542
7	lying down	0.2379
8	standing up from lying	0.1086
9	sit to stand	0.4019
10	stand to sit	0.2582
11	sitting still	0.6957
12	jumping slow	0.9487

Table III
ACCURACY OF THE CLASSIFIER ON THE MERGED VERSION OF THE DATABASE.

Table IV shows that the simple union of the databases, with the exception of only two cases (3 and 20 respectively), leads to an overall worsening of the accuracy compared to the use of the single database (*Original* column) on average equal to -0.18. This confirms the fact that it is not possible to simply unify the databases without first analysing the labels.

By applying our procedure instead, it is possible to notice a small worsening compared to the use of the single database (-0.14 on average). However, in 10 cases out of 24 (42% of cases) there is an improvement equals to 0.11 on average. The result obtained is to be considered very positive because the unified data are very heterogeneous being acquired by different sets of people who carried out activities in different contexts and with different acquisition procedures.

V. CONCLUSIONS

In this paper we present a semi-automatic procedure to merge multiple databases of inertial signals employed by the machine learning research community for Human Activity Recognition. The purpose of this procedure is to cope with the incessant growth of available labeled data that is due

Number	Labels	Original	Union	Merging
1	walking UniMiB	0.8466	0.4842	0.8168
2	running UniMiB	0.8977	0.7570	0.8705
3	going upstairs UniMiB	0.2582	0.3040	0.6012
4	going downstairs UniMiB	0.3026	0.1235	0.5059
5	jumping UniMiB	0.5949	0.3912	0.3542
6	lying down fs UniMiB	0.3943	0.2467	0.2379
7	standing up fs UniMiB	0.6263	0.3232	0.3826
8	standing up fl UniMiB	0.1858	0.0849	0.1086
9	sitting down UniMiB	0.4437	0.2932	0.2911
10	walking MotionSense	0.9038	0.7772	0.9837
11	jogging MotionSense	0.9277	0.6545	0.9921
12	upstairs MotionSense	0.6728	0.2103	0.4159
13	downstairs MotionSense	0.3417	0.1608	0.3903
14	standing MotionSense	0.4419	0.2536	0.6304
15	sitting MotionSense	0.8051	0.7066	0.6957
16	walking Mobiact	0.9851	0.9569	0.9913
17	jogging Mobiact	0.9328	0.8752	0.9444
18	stairs up Mobiact	0.7038	0.6251	0.6966
19	stairs down Mobiact	0.5181	0.4391	0.6397
20	jumping Mobiact	0.9780	0.9780	0.9487
21	standing Mobiact	0.9607	0.4669	0.6408
22	sit to stand Mobiact	0.3791	0.2330	0.4019
23	sitting on chair Mobiact	0.3120	0.2551	0.1935
24	stand to sit Mobiact	0.4859	0.4523	0.2582

Table IV
ACCURACY OF THE CLASSIFIER ON THE SINGLE DATABASE, ON A UNIFIED AND MERGED VERSION OF THE DATABASE.

to the increasingly use of wearable acquisition devices in daily life. Labeled databases are usually not consistent, both syntactically (e.g., different sampling frequency) and semantically (e.g., labels with different meanings). This makes it very difficult for machine learning researchers to take full advantage of having so much data available. The proposed procedure performs a syntactic and semantic merging of labels belonging to several database by using signal and word similarity.

Preliminary results show the effectiveness of the proposed approach. In particular, classification experiments, performed using a k-NN algorithm, show that a classifier trained on a merged database is capable of, in most of the cases, discriminating labels much better than the case where the labels are not merged.

Although the results achieved are promising, this is only a preliminary study. A further step should be done in order to experiment the proposed procedure on a very large set of databases. Most important, classification should be performed using deep learning that are machine learning approach more sensitive to the size of the training set.

Moreover, when the procedure will be accurately tested, it can be integrated in the Continuous Learning Platform that requires an homogenization module in order to integrate heterogeneous databases and distribute them in an homogeneous form [38].

REFERENCES

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192–1209, 2013.

[2] D. Micucci, M. Mobilio, P. Napoletano, and F. Tisato, "Falls as anomalies? an experimental evaluation using smartphone accelerometer data," *Journal of Ambient Intelligence and Humanized Computing*, vol. 8, no. 1, pp. 87–99, 2017.

[3] H.-B. Zhang, Y.-X. Zhang, B. Zhong, Q. Lei, L. Yang, J.-X. Du, and D.-S. Chen, "A comprehensive survey of vision-based human action recognition methods," *Sensors*, vol. 19, no. 5, 2019.

[4] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, p. 436, 2015.

[5] G. Hu, Y. Yang, D. Yi, J. Kittler, W. Christmas, S. Z. Li, and T. Hospedales, "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition," in *Proceedings of the 2015 IEEE International Conference on Computer Vision Workshop (ICCVW15)*, 2015.

[6] J. Bartlett, V. Prabhu, and J. Whaley, "Acctionnet: A manusing on-phone motion sensors," in *Proceedings of the 34th International Conference on Machine Learning (ICML17)*, 2017.

[7] D. Ginelli, D. Micucci, M. Mobilio, and P. Napoletano, "UniMiB AAL: An Android Sensor Data Acquisition and Labeling Suite," *Applied Sciences*, vol. 8, no. 8, 2018.

[8] D. Dua and C. Graff, "UCI machine learning repository." http://archive.ics.uci.edu/ml, 2017. Online; accessed 29 March 2019.

[9] P. Siirtola, H. Koskimäki, and J. Röning, "OpenHAR: A Matlab Toolbox for Easy Access to Publicly Open Human Activity Data Sets," in *Proceedings of the ACM International Joint Conference and International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers (UbiComp18)*, 2018.

[10] S. Katz, T. D. Downs, H. R. Cash, and R. C. Grotz, "Progress in development of the index of adl," *The gerontologist*, vol. 10, no. 1_Part_1, pp. 20–30, 1970.

[11] L. S. Noelker and R. Browdie, "Sidney Katz, MD: A new paradigm for chronic illness and long-term care," *The Gerontologist*, vol. 54, no. 1, pp. 13–20, 2014.

[12] J. M. Wiener, R. J. Hanley, R. Clark, and J. F. Van Nostrand, "Measuring the activities of daily living: Comparisons across national surveys," *Journal of gerontology*, vol. 45, no. 6, pp. S229–S237, 1990.

[13] W. D. Spector, S. Katz, J. B. Murphy, and J. P. Fulton, "The hierarchical relationship between activities of daily living and instrumental activities of daily living," *Journal of chronic diseases*, vol. 40, no. 6, pp. 481–489, 1987.

[14] M. P. Lawton and E. M. Brody, "Assessment of older people: self-maintaining and instrumental activities of daily living," *The Gerontologist*, vol. 9, no. 3,Pt1, pp. 179–186, 1969.

[15] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "A survey of online activity recognition using mobile phones," *Sensors*, vol. 15, no. 1, pp. 2059–2085, 2015.

[16] O. D. Lara, A. J. Perez, M. A. Labrador, and J. D. Posada, "Centinela: A human activity recognition system based on acceleration and vital sign data," *Pervasive Mob. Comput.*, vol. 8, no. 5, pp. 717–729, 2012.

[17] M. Shoaib, S. Bosch, O. Incel, H. Scholten, and P. Havinga, "Fusion of smartphone motion sensors for physical activity recognition," *Sensors*, vol. 14, no. 6, pp. 10146–10176, 2014.

[18] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones.," in *Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, (ESANN13)*, 2013.

[19] J. J. Guiry, P. van de Ven, J. Nelson, L. Warmerdam, and H. Riper, "Activity recognition with smartphone support," *Medical engineering & physics*, vol. 36, no. 6, pp. 670–675, 2014.

[20] D. Anguita, A. Ghio, L. Oneto, F. X. Llanas Parra, and J. L. Reyes Ortiz, "Energy efficient smartphone-based activity recognition using fixed-point arithmetic," *Journal of universal computer science*, vol. 19, no. 9, pp. 1295–1314, 2013.

[21] Q. Li, J. A. Stankovic, M. A. Hanson, A. T. Barth, J. Lach, and G. Zhou, "Accurate, fast fall detection using gyroscopes and accelerometer-derived posture information.," in *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks (BSN09)*, 2009.

[22] X. Su, H. Tong, and P. Ji, "Activity recognition with smartphone sensors," *Tsinghua science and technology*, vol. 19, no. 3, pp. 235–249, 2014.

[23] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, "Simple and complex activity recognition through smart phones," in *Proceedings of the International Conference on Intelligent Environments (IE12)*, 2012.

[24] J.-H. Hong, J. Ramos, and A. K. Dey, "Toward personalized activity recognition systems with a semipopulation approach," *IEEE Transactions on Human-Machine Systems*, vol. 46, no. 1, pp. 101–112, 2016.

[25] M. Zhang and A. A. Sawchuk, "A feature selection-based framework for human activity recognition using wearable multimodal sensors," in *Proceedings of the International Conference on Body Area Networks (BODYNETS 11)*, 2011.

[26] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," *Sensors*, vol. 15, no. 12, pp. 31314–31338, 2015.

[27] M. Kose, O. D. Incel, and C. Ersoy, "Online human activity recognition on smart phones," in *Proceedings of the Workshop on Mobile Sensing: From Smartphones and Wearables to Big Data*, 2012.

[28] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SIGKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[29] Z. He and L. Jin, "Activity recognition from acceleration data based on discrete consine transform and svm," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC09)*, 2009.

[30] Y. Vaizman, N. Weibel, and G. Lanckriet, "Context recognition in-the-wild: Unified model for multi-modal sensors and multi-label classification," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT)*, vol. 1, no. 4, pp. 168:1–168:22, 2018.

[31] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proceedings of the Conference on Innovative applications of Artificial Intelligence (IAAI 05)*, 2005.

[32] D. Micucci, M. Mobilio, and P. Napoletano, "UniMiB SHAR: A dataset for human activity recognition using acceleration data from smartphones," *Applied Sciences*, vol. 7, no. 10, p. 1101, 2017.

[33] S. Bianco, P. Napoletano, and R. Schettini, "Multimodal car driver stress recognition," in *Proceedings of the Workshop on Affective Computing in Pervasive Health (AffectPH19)*, 2019. in press.

[34] G. Vavoulas, C. Chatzaki, T. Malliotakis, M. Pediaditis, and M. Tsiknakis, "The mobiact dataset: Recognition of activities of daily living using smartphones.," in *Proceedings of Information and Communication Technologies for Ageing Well and e-Health (ICT4AgeingWell16)*, 2016.

[35] M. Malekzadeh, R. G. Clegg, A. Cavallaro, and H. Haddadi, "Protecting sensory data against sensitive inferences," in *Proceedings of the Workshop on Privacy by Design in Distributed Systems (W-P2DS18)*, 2018.

[36] T. Sztyler, J. Carmona, J. Völker, and H. Stuckenschmidt, "Self-tracking reloaded: Applying process mining to personalized health care from labeled sensor data," *Transactions on Petri Nets and Other Models of Concurrency XI*, vol. 9930, pp. 160–180, 2016.

[37] E. Casilari, J. A. Santoyo-Ramón, and J. M. Cano-García, "Umafall: A multisensor dataset for the research on automatic fall detection," *Procedia Computer Science*, vol. 110, pp. 32–39, 2017.

[38] A. Ferrari, D. Micucci, M. Marco, and P. Napoletano, "A framework for long-term data collection to support automatic human activity recognition," in *Proceedings of Intelligent Environments: Workshop on Reliable Intelligent Environment (IE 19)*, 2019.