Robust Human Activity Recognition Using Lesser Number of Wearable Sensors

Di Wang[†], Edwin Candinegara[‡], Junhui Hou[§], Ah-Hwee Tan[†‡] and Chunyan Miao[†‡]
[†]Joint NTU-UBC Research Center of Excellence in Active Living for the Elderly
[‡]School of Computer Science and Engineering
Nanyang Technological University
Singapore
[§]Department of Computer Science
City University of Hong Kong
Hong Kong
wangdi@ntu.edu.sg, edwin.candinegara@gmail.com, jh.hou@cityu.edu.hk, asahtan@ntu.edu.sg, ascymiao@ntu.edu.sg

***Abstract*—In recent years, research on the recognition of human physical activities solely using wearable sensors has received more and more attention. Compared to other types of sensory devices such as surveillance cameras, wearable sensors are preferred in most activity recognition applications mainly due to their non-intrusiveness and pervasiveness. However, many existing activity recognition applications or experiments using wearable sensors were conducted in the confined laboratory settings using specifically developed gadgets. These gadgets may be useful for a small group of people in certain specific scenarios, but probably will not gain their popularity because they introduce additional costs and they are unusual in everyday life. Alternatively, commercial devices such as smart phones and smart watches can be better utilized for robust activity recognitions. However, only few prior studies focused on activity recognitions using multiple commercial devices. In this paper, we present our feature extraction strategy and compare the performance of our feature set against other feature sets using the same classifiers. We conduct various experiments on a subset of a public dataset named PAMAP2. Specifically, we only select two sensors out of the thirteen used in PAMAP2. Experimental results show that our feature extraction strategy performs better than the others. This paper provides the necessary foundation towards robust activity recognition using only the commercial wearable devices.**

***Index Terms*—activity recognition, PAMAP2 dataset, wearable sensor, support vector machine, random forest**

I. Introduction

Human physical activity recognition is a challenging but emerging research topic. In essence, a system is built to recognize the specific types of activities performed by the human subjects using data collected from sensors. Many application areas benefit from a robust activity recognition system, such as healthcare, security, sports, etc. In terms of healthcare for the general population, activity tracker mobile devices and apps are pervasive nowadays. Moreover, activity recognition may be much more beneficial to the elderly to query their well-being, ensure their safety, provide personalized recommendations on healthier activities, etc.

This research is supported in part by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative.

For activity recognition within a predefined area, surveillance cameras, such as video cameras, infra-red cameras, motion detectors, are often used, which are more and more reliable due to the recent advances in image and video understanding using machine learning techniques [1], [2]. However, such surveillance systems always receive privacy concerns due to their intrusiveness as people feel being watched.

Another dominating type of activity recognition approach is to use the sensory inputs collected from wearable sensors or devices. However, many such applications or experiments were conducted in the confined laboratory settings using specifically developed gadgets [3], [4]. These gadgets may be useful for a small group of people in certain specific scenarios, but probably will not gain their popularity because they introduce additional costs and they are unusual in everyday life.

On the other hand, commercial devices such as smart phones and smart watches are pervasive in most developed and developing countries that "more than half the world now uses a smart phone" [5]. However, the types of activities recognizable by the commercial devices are still limited. Smart wristbands and watches can count the number of your steps on a daily basis and some can log your sleep status. Smart phones can recognize and record much more types of activities, e.g., Android smart phones may recognize 'In vehicle', 'On bicycle', 'On foot', 'Running', 'Still', 'Tilting' and 'Walking', using the ActivityRecognitionApi [6]. Nonetheless, all these types of activities are primitive that they do not even cover the basic Activities of Daily Living (ADLs), not to say instrumental ADLs [7]. Therefore, the motivation of our work is as follows:

> *Whether we can use a limited number of commercial devices to accurately recognize a large variety of physical activities solely based on the signals collected from the sensors embedded in those devices?*

Towards using lesser number of sensors embedded in the commercial wearable devices, an activity recognition system faces two major challenges to achieve satisfactory robustness, namely the varying positions of the devices and the confidence in recognizing the same complex activities performed in various patterns due to individual differences.

The challenge of varying positions of the devices is not the focus of this paper. Readers may refer to other publications on robust activity recognition using a single device (smart phone) with varying orientations, facing and placement [8], [9]. Readers may also refer to other publications on using a single device (smart phone) for both activity recognition and context awareness based on the placement of the phone and the surrounding environment [10]–[12].

Alternatively, this paper focuses on meeting the challenge of activity recognition with individual different behavioral patterns. Specifically, we design and conduct our study according to the following three criteria: (i) the set of activities to be recognized should cover more than the primitive activities, (ii) activity recognition should be performed using lesser number of sensors yet the recognition should achieve a reasonable level of accuracy, and (iii) most importantly, the activity recognition algorithm should perform robustly even if the underlying behavioral patterns are unseen.

To test whether our extracted feature set, which consists of features extracted from both the time and frequency domains (see Section III), can meet the challenges of using lesser number of sensors but still performing well even on unseen behavioral patterns, we conducted extensive experiments on a public dataset named PAMAP2 [13] [14], which collected the raw sensory input data of twelve types of activities performed by nine subjects. However, we did not use all the sensory readings archived in PAMAP2. Instead, we only selected two sensors out of the total number of thirteen (see Section IV-A). Furthermore, we conducted experiments on applying other feature sets [15] [16] on the same selected dataset for performance comparisons. Experimental results show that our feature set outperforms the other two in the leave-one-person-out scenario. Moreover, our feature set obtains satisfactory level of performance when compared with the results reported in [13] and [14], if taking the number of sensors in use, window size used for recognition, sampling rate, and the number of extracted features into consideration. In summary, our feature set is empirically shown to be accurate in activity recognitions using lesser number of sensors and robust in dealing with unseen individual different behavioral patterns.

The rest of the paper is organized as follows. In Section II, we review the related work. In Section III, we introduce our model for robust human activity recognition. In Section IV, we present the experimental results of our activity recognition model with comparisons and discussions. In Section V, we conclude this paper and propose future extensions.

II. Related Work

To maximize activity recognition accuracy, some researchers opt to design and implement their own devices or gadgets to monitor the body movements precisely. Nam and Park [3] developed a wearable device that consists of several sensors (including accelerometer and barometer) to recognize eleven types of activities of ten children from 16 to 29 months of age. Reiss and Stricker [13] [14] used three inertial measurement units (IMUs) and a heart rate monitor to recognize twelve types of activities of nine subjects aged 27 in average. Chernbumroong et al. [4] developed a set of wearable devices (one chest strap for heart rate and two wrist watches with integrated accelerometer, altimeter and other sensors) to recognize twelve types of activities of the elderly aged 73 in average. All these studies show promising results. However, the usage of these specially crafted or not-everyday-in-use devices may be hindered due to the popularity reason. We prefer to use widely owned commercial devices for non-intrusive human activity recognition.

Studies on activity recognition using commercial devices have been emerging in recent years, due to the ever increasing penetration rate of those devices. However, majority studies only used one device, which may not cover a comprehensive set of activities to be recognized. Some pioneer studies (e.g., [17] and [18]) only used the accelerometer readings for activity recognition. Later on, more studies included other sensory inputs from the same device to improve the activity recognition accuracy, such as gyroscope [19], light and proximity sensors [12], Wi-Fi and GPS signals [20], etc. Although we prefer to use lesser number of sensors, we still want to use sensors placed in different locations, e.g., movement tracking of the wrist may allow us to distinguish writing and typing.

Only few prior studies focused on activity recognitions using multiple commercial devices that the most common combination is one smart watch placed on the wrist (dominant hand) and one smart phone placed in a fixed place such as attached to the belt or in a pocket. Nonetheless, although using more than one devices, some studies still focused on a limited number of primitive activities (e.g., five in [21] and nine in [22]). Although Shoaib et al. [15] mimicked the deployment of a smart watch by attaching a smart phone to the wrist, they achieved quite good recognition accuracy on thirteen activities including those only distinguishable by wrist movements such as typing, writing, eating, drinking coffee, and smoking. The activities studied in [16] are richer and more challenging, such as clapping hands, brushing teeth, folding clothes, and eating various kinds of food. The complex activities studied in [15] and [16] allow us to learn more of one's daily activities non-intrusively. However, datasets used in [15] and [16] are not publicly available. Therefore, in this paper, we use the PAMAP2 dataset instead. In PAMAP2, the IMUs were placed on the subject's wrist, chest and ankle. We select one accelerometer on the wrist by assuming it is embedded in a smart watch and another one on the chest by assuming it is embedded in a smart phone being placed in one of the jacket's pockets. Nonetheless, both the feature sets used in [15] and [16] are also used in this paper for benchmarks.

III. Robust Activity Recognition Model

Because every smart phone and smart watch has an embedded triaxial accelerometer, our robust activity recognition (**RAR**) model consists of the following eight stages: (i) collect the raw sensory readings from each accelerometer,

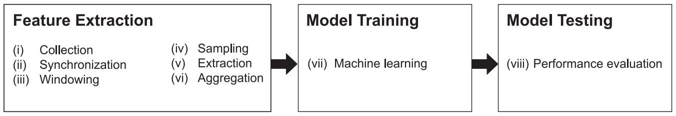

Fig. 1. The overall work flow of our robust activity recognition model.

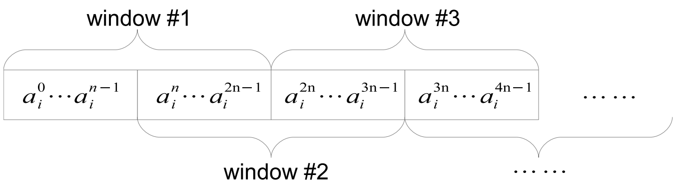

Fig. 2. Formation of half overlapping windows ($i \in \{x, y, z, mag\}$, where mag denotes the acceleration magnitude, see (2)). In this illustration, the frequency is n Hz and window size is 2 seconds.

(ii) synchronize the various signals, (iii) organize them into data windows, (iv) sample them using the same frequency, (v) extract both the time and frequency domain features, (vi) aggregate the features from different sources to obtain the overall feature set, (vii) apply machine learning algorithms on the overall feature set to distinguish various types of activities, and (viii) evaluate the performance of the recognition. For simplicity, stages (i)-(vi) can be grouped as the feature extraction stage. The overall work flow of our robust activity recognition model is shown in Fig. 1. In the following subsections, we introduce the feature extraction strategies, machine learning algorithms, and performance evaluation metrics, respectively.

A. Feature Extraction Strategies

There are three readings ($A = \{a_x, a_y, a_z\}$) given by any triaxial accelerometer at any time (t) corresponding to the three orthogonal axes (x, y, z), respectively:

$$A^t = \{a_x^t, a_y^t, a_z^t\}. \tag{1}$$

Because we only use accelerometers, which do not directly tell any orientation information. For robust activity recognition without knowing the orientation of the device (e.g., from gyroscope or magnetometer), we introduce an extended dimension (acceleration magnitude) as augmentation [23] [9]. Therefore, at any time (t), we use four readings that are defined as

$$A^t = \{a_x^t, a_y^t, a_z^t, a_{mag}^t\}, \tag{2}$$

where mag denotes the acceleration magnitude ($||x, y, z||$) and $||x, y, z|| = \sqrt{(x^2 + y^2 + z^2)}$.

There are two parameters, window size and frequency, to be defined for data formation. After which, continuous data will be chopped into discrete windows (see Fig. 2). In this paper, all windows are half overlapping.

After data formation, assume each window consists of N readings. Then, in each dimension ($i \in \{x, y, z, mag\}$), we extract four features, namely mean (M), variance (V), energy (EE), and entropy (ET). The first two features are defined as

$$M_i = \frac{1}{N}\sum_{n=0}^{N-1} a_i^n \tag{3}$$

and

$$V_i = \frac{1}{N-1}\sum_{n=0}^{N-1}(a_i^n - M_i)^2, \tag{4}$$

respectively. The latter two are computed in the frequency domain, where discrete Fourier transform (DFT) is applied:

$$F_i(k) = \sum_{n=0}^{N-1} a_i^n e^{-j2\pi nk/N}, k = 0, 1, 2, \ldots, N-1. \tag{5}$$

The energy ($L2$ norm) and entropy are then defined as

$$EE_i = \sqrt{\frac{\sum_{k=1}^{N-1}|F_i(k)|^2}{N-1}} \tag{6}$$

and

$$ET_i = \sum_{l=1}^{N-1} -O_i(l)\ln(O_i(l)), \tag{7}$$

where

$$O_i(l) = \frac{|F_i(l)|}{\sum_{k=1}^{N-1}|F_i(k)|}, \tag{8}$$

respectively. Please note that since the DC component of (5) (mean) has already been used as an individual feature (see (3)), when computing the energy and entropy, the indices start from 1 instead of 0.

Up to now, all the extracted features are from the individual axes. To make the activity recognition more robust, covariances (COV) between the axes (of each accelerometer embedded in different devices) are also used:

$$COV_{p,q} = \frac{1}{N-1}\sum_{t=0}^{N-1}(a_p^t - M_p)(a_q^t - M_q), \tag{9}$$

where $p, q \in \{x, y, z, mag\}$ and $p \neq q$.

Therefore, after introducing the six covariance measures, the total number of features extracted from each triaxial accelerometer is $4*4+6 = 22$. Furthermore, after aggregation of two accelerometers embedded in two different devices, the total number of features in the overall feature set is 44.

Normalization is performed on all the extracted feature values before being processed by any machine learning algorithm. Let f denote the index of the feature ($f = 1, 2, \ldots, 44$), then the maximum and minimum values of all the observed data in each feature can be denoted as $\max_f$ and $\min_f$, respectively. Therefore, all values (both observed and unobserved) will be normalized using the following equation:

$$v'_f = \frac{v_f - \min_f}{\max_f - \min_f}, \tag{10}$$

where v'_f denotes the normalized value in the fth feature and v_f denotes the original value in the fth feature.

B. Machine Learning Algorithms

In this paper, we use two well-known machine learning algorithms as the classifiers for activity recognition, namely random forest and support vector machine (SVM). Furthermore, we use scikit-learn [24] to implement both algorithms.

Random forest [25] is an ensemble algorithm. As its name suggests, a random forest model consists of a multitude of decision trees, where each decision tree is trained on a randomly selected subset of the training data and the feature sets. During prediction, the answers produced by each decision tree are aggregated by various means of majority voting. In scikit-learn, a random forest model derives the final prediction using probabilistic prediction, which is defined as follows:

$$A = \underset{c}{\operatorname{argmax}} \frac{1}{N} \sum_{i=1}^{N} P_i(c), \tag{11}$$

where A denotes the predicted class, c denotes the class label, N denotes the number of decision trees, and $P_i(c)$ denotes the probability of class c computed by the ith decision tree.

SVM [26] aims to minimize the structural risk of the learnt model, especially when there are a limited number of training data available. In this paper, we use the radial basis function (RBF) kernel to transform the training data into a higher dimensional space. For SVM with RBF kernel, there are two control parameters to be defined. One is the cost parameter C ($C > 0$), which defines the amount of penalty on error. The other is the gamma parameter γ ($\gamma > 0$), which controls the flexibility of the decision boundary. Both C and γ control the level of generalization of the SVM model. Their optimal values should prevent both the over- and less-fitting problems. In this paper, we set the values of C and γ in two ways. If C and γ are determined by a grid search using cross-validation, we denote the learnt model as **SVM-best**. If the default values, i.e., $C = 1$ and $\gamma = 1/l$, where l denotes the dimensionality of the feature space, are used as suggested [27], we denote the learnt model as **SVM-default**.

C. Performance Evaluation Metrics

After training, the learnt models are applied to the testing dataset for performance evaluation. To measure both Type-I and Type-II errors, precision and recall are defined as

$$\text{precision} = \frac{\sum \text{true positive}}{\sum \text{predicted positive}} = \frac{TP}{TP + FP} \tag{12}$$

and

$$\text{recall} = \frac{\sum \text{true positive}}{\sum \text{actual positive}} = \frac{TP}{TP + FN}, \tag{13}$$

respectively. The terms TP, FP, and FN denote true positive, false positive, and false negative, respectively.

The overall performance (F-score) is then computed as

$$F\text{-score} = \frac{2 * \text{precision} * \text{recall}}{\text{precision} + \text{recall}}. \tag{14}$$

IV. Experimental Results and Analysis

In this section, we present the details of how we conduct the experiments to evaluate our approach of using lesser number of sensors for robust recognition of a wide range of activities. Furthermore, we compare the experimental results with other two approaches and provide the corresponding analysis.

A. Dataset Description

The PAMAP2 dataset [13] [14] collected the raw movement sensory data of twelve different activities performed by nine subjects aged 27.22 ± 3.31 in average. The twelve activities are (i) lying, (ii) sitting, (iii) standing, (iv) walking, (v) running, (vi) cycling, (vii) nordic walking, (viii) going upstairs, (ix) going downstairs, (x) vacuum cleaning, (xi) ironing, and (xii) rope jumping. The aggregated time of all activities performed by all subjects is approximately 5.5 hours.

There were four gadgets used during data collection, namely three inertial measurement units (IMUs) and one heart rate monitor. The three IMUs were placed on the subject's wrist (dominant hand), chest and ankle, respectively. The heart rate monitor was placed on the chest to collect heart beat signals. Each IMU integrates four sensors, namely two triaxial accelerometers (± 16 g and ± 6 g), one gyroscope and one magnetometer. Therefore, there are a total number of thirteen sensors used in PAMAP2.

Out of the thirteen sensors, in our experiments, we only select two accelerometers that were placed on the wrist and chest, respectively, because they can be considered as mimicking the accelerometers embedded in a smart watch being placed on the wrist and a smart phone being placed in one of the jacket's pockets. The two selected accelerometers are of ± 6 g, which is closer to the specifications of those embedded in smart watches and phones. Furthermore, to make our experiments more towards real usages in one's daily life rather than in the laboratory setting, we greatly downsample the sensory readings (from 100 Hz to 5 Hz, for energy conservation) and greatly reduce the window size (from 5.12 s to 2 s, for faster and more challenging recognition).

B. Benchmarking Feature Sets

We denote the set of features extracted using Shoaib et al.'s approach [15] as Shoaib's feature set. This feature set consists of 28 extracted features per sensor. Specifically, seven features, namely mean, standard deviation, min, max, median, semi-quartile, and the sum of the first ten DFT coefficients, are extracted from each axis (x, y, z and mag).

We denote the set of features extracted using Weiss et al.'s approach [16] as Weiss's feature set. This feature set consists of 43 extracted features per sensor. Specifically, fourteen features, namely mean, standard deviation, average absolute difference, time between peaks, and binned distribution (ten bins per axis), are extracted from each axis (x, y and z). In addition, the last feature is the average resultant acceleration. Please note that the "time between peaks" features (for x, y and z) are excluded in this paper, because they require certain heuristic assessments that were not described in [16].

TABLE I
PERFORMANCE EVALUATIONS ON 5-FOLD CROSS VALIDATION

Feature Set	Random Forest	SVM-best	SVM-default
RAR	0.9458	0.9476	0.8142
Shoaib [15]	**0.9584**	0.9567	0.7371
Weiss [16]	0.9344	0.8948	0.7469

TABLE II
PERFORMANCE EVALUATIONS ON LEAVE-ONE-PERSON-OUT

Feature Set	Random Forest	SVM-default
RAR	**0.8062**	0.7216
Shoaib [15]	0.7829	0.6537
Weiss [16]	0.7591	0.6074

C. Experimental Setups

In this paper, we use three types of classifiers, namely random forest, SVM-best and SVM-default. For random forest, we always employ 500 decision trees. For SVM-best, the optimal values of C and γ are determined by a grid search on the training dataset. The range of the grid search is $C \in \{1, 10, 100, 500, 1000, 5000\}$ and $\gamma \in \{1/l, 0.01, 0.05, 0.1, 0.5, 1\}$, respectively. For SVM-default, the values for C and γ are always set to 1 and $1/l$, respectively. All experimental results shown in this paper are the average of three runs aiming to remove the effect of randomness.

We evaluate the performance of all feature sets in two scenarios: (i) 5-fold cross validation (CV), wherein each time four fifths of all subjects' data are used for training and the remaining one fifth are used for testing, and (ii) leave-one-person-out (LOPO), wherein each time eight out of the nine subjects' data are used for training and the remaining one subject's data are used for testing. For 5-fold CV, we report the results of all the three classifiers. However, for LOPO, we do not report the results of SVM-best, because it always performs significantly worse than the other two classifiers, probably due to over-fitting on the observed behavioral patterns.

D. Experimental Results

The experimental results of applying three classifiers on all the three feature sets in the 5-fold CV scenario is shown in Table I. As shown in Table I, the combination of applying random forest on Shoaib's feature set achieves the best F-score. However, the difference of the best performance achieved by our RAR feature set (applying SVM-best) is as small as $0.9584 - 0.9476 = 0.0108$ or 1.08%. The performance of all feature sets are generally good as all of them can achieve an F-score greater than 0.93 in distinguishing twelve activities.

The experimental results of applying two classifiers on all the three feature sets in the LOPO scenario is shown in Table II. It is encouraging to find out that our RAR feature set achieves the best performance in recognizing the activities performed by a subject whose behavioral patterns are not learned before. The difference between RAR and the runner-up (Shoaib's feature set) is $0.8062 - 0.7829 = 0.0233$ or 2.33%.

TABLE III
SIZE OF ALL THE FEATURE SETS

Feature Set	RAR	Shoaib [15]	Weiss [16]
# features	**44**	56	80

TABLE IV
F-SCORES ON LOPO USING DIFFERENT NUMBER OF SENSORS

# sensors	# features	window size	frequency	Classifier	F-score
2	**44**	**2 seconds**	**5 Hz**	Random Forest	0.8062
13	137	5.12 seconds	100 Hz	C4.5	0.8300
13	137	5.12 seconds	100 Hz	Boosted C4.5	0.8928
13	137	5.12 seconds	100 Hz	Bagging C4.5	0.8556
13	137	5.12 seconds	100 Hz	Naive Bayes	0.8362
13	137	5.12 seconds	100 Hz	kNN	**0.9110**

The F-scores of the bottom five classifiers are taken from [13] and [14].

As aforementioned in Section I, the aim of this research is to build a robust activity recognition system using lesser number of sensors to distinguish complex activities with various patterns (especially unobserved) due to individual differences. In view of this, the performance of the feature sets is mainly defined by their accuracy in the LOPO scenario. Moreover, considering RAR achieves the best LOPO performance using the least number of features (see Table III), we may say that our RAR feature set outperforms the other two features sets.

E. Analysis on Using Lesser Number of Sensors

The performance comparison between the best performance obtained using two sensors and that using thirteen sensors is presented in Table IV. There seems to be some difference between the performance of our RAR feature set with that of the others. However, if you take the amount of information used in the whole activity recognition process into consideration, i.e., number of sensors in use (2 vs 13), number of extracted features (44 vs 137), window size (2 s vs 5.12 s), and sampling frequency (5 Hz vs 100 Hz), our approach is definitely much more practical in real-world settings.

V. CONCLUSION

This paper presents our feature extraction strategy towards robust activity recognition using lesser number of wearable sensors embedded in commercial devices. Moreover, the performance of our extracted feature set is evaluated using a public dataset and compared with two recently published, state-of-the-art benchmarking feature sets. The experimental results show that our feature set performs competitively in the cross-validation scenario and more importantly, it outperforms the benchmarking feature sets in the leave-one-person-out scenario. Furthermore, even our approach uses significantly lesser amount of information across the whole activity recognition process, its performance is still satisfactory.

Going forward, we plan to extend our work in the following two aspects. First, we will try to include more informative and robust features into our feature set to boost up the accuracy. Secondly, we will recruit subjects and collect their activity data using only the commercial-off-the-shelf smart watches and phones in daily life settings.

References

[1] Z. He, X. Li, D. Tao, X. You, and Y. Y. Tang, "Connected component model for multi-object tracking," *IEEE Transactions on Image Processing*, vol. 25, no. 8, pp. 3698–3711, 2016.

[2] Z. He, S. Yi, Y.-M. Cheung, X. You, and Y. Y. Tang, "Robust object tracking via key patch sparse representation," *IEEE Transactions on Cybernetics*, vol. 47, pp. 354–364, 2017.

[3] Y. Nam and J. W. Park, "Child activity recognition based on cooperative fusion model of a triaxial accelerometer and a barometric pressure sensor," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 2, pp. 420–426, 2013.

[4] S. Chernbumroong, S. Cang, and H. Yu, "A practical multi-sensor activity recognition system for home-based care," *Decision Support Systems*, vol. 66, pp. 61–70, 2014.

[5] We Are Social. (2017, January) Digital in 2017: Global overview. [Online]. Available: https://wearesocial.com/sg/blog/2017/01/digital-in-2017-global-overview/

[6] Google APIs for Android, "ActivityRecognitionApi." [Online]. Available: https://developers.google.com/android/reference/com/google/android/gms/location/ActivityRecognitionApi/

[7] S. Gao, D. Wang, A.-H. Tan, and C. Miao, "Progressive sequence matching for ADL plan recommendation," in *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. IEEE, 2015, pp. 360–367.

[8] L. Sun, D. Zhang, and N. Li, "Physical activity monitoring with mobile phones," in *Toward Useful Services for Elderly and People with Disabilities*, ser. Lecture Notes in Computer Science, B. Abdulrazak, S. Giroux, B. Bouchard, H. Pigot, and M. Mokhtari, Eds. Springer, 2011, vol. 6719, pp. 104–111.

[9] D. Wang, A.-H. Tan, and D. Zhang, "Non-intrusive robust human activity recognition for diverse age groups," in *Proceedings of IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*. IEEE, 2015, pp. 368–375.

[10] P. Korpipaa, J. Mantyjarvi, J. Kela, H. Keranen, and E.-J. Malm, "Managing context information in mobile devices," *Pervasive Computing*, vol. 2, no. 3, pp. 42–51, 2003.

[11] E. Miluzzo, C. T. Cornelius, A. Ramaswamy, T. Choudhury, Z. Liu, and A. T. Campbell, "Darwin phones: The evolution of sensing and inference on mobile phones," in *Proceedings of International Conference on Mobile Systems, Applications, and Services*, 2010, pp. 5–20.

[12] H. Martin, A. M. Bernardos, J. Iglesias, and J. R. Casar, "Activity logging using lightweight classification techniques in mobile devices," *Personal and Ubiquitous Computing*, vol. 17, no. 4, pp. 675–695, 2013.

[13] A. Reiss and D. Stricker, "Introducing a new benchmarked dataset for activity monitoring," in *Proceedings of IEEE International Symposium on Wearable Computers*, 2012, pp. 108–109.

[14] ——, "Creating and benchmarking a new dataset for physical activity monitoring," in *Proceedings of Workshop on Affect and Behaviour Related Assistance*, no. 40, 2012.

[15] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. Havinga, "Complex human activity recognition using smartphone and wrist-worn motion sensors," *Sensors*, vol. 16, no. 4, p. 426, 2016.

[16] G. M. Weiss, J. L. Timko, C. M. Gallagher, K. Yoneda, and A. J. Schreiber, "Smartwatch-based activity recognition: A machine learning approach," in *Proceedings of IEEE-EMBS International Conference on Biomedical and Health Informatics*, 2016, pp. 426–429.

[17] J. Baek, S. T. Kim, H. D. Kim, J.-S. Cho, and B.-J. Yun, "Recognition of user activity for user interface on a mobile device," in *Proceedings of South East Asia Regional Computer Conference*, 2007, pp. 10.1–10.6.

[18] J. Yang, "Toward physical activity diary: Motion recognition using simple acceleration features with mobile phones," in *Proceedings of International Workshop on Interactive Multimedia for Consumer Electronics*, 2009, pp. 1–10.

[19] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine," in *Ambient Assisted Living and Home Care*, ser. Lecture Notes in Computer Science, J. Bravo, R. Hervas, and M. Rodriguez, Eds. Springer, 2012, vol. 7657, pp. 216–223.

[20] L. Pei, R. Guinness, R. Chen, J. Liu, H. Kuusniemi, Y. Chen, L. Chen, and J. Kaistinen, "Human behavior cognition using smartphone sensors," *Sensors*, vol. 13, no. 2, pp. 1402–1424, 2013.

[21] F. Ramos, A. Moreira, A. Costa, R. Rolim, H. Almeida, and A. Perkusich, "Combining smartphone and smartwatch sensor data in activity recognition approaches: An experimental evaluation," in *Proceedings of International Conference on Software Engineering and Knowledge Engineering*, 2016, pp. 267–272.

[22] A. Dengel, B. Sefen, S. Baumbach, and S. Abdennadher, "Human activity recognition: Using sensor data of smartphones and smartwatches," in *Proceedings of International Conference on Agents and Artificial Intelligence*, 2016, pp. 488–493.

[23] L. Sun, D. Zhang, B. Li, B. Guo, and S. Li, "Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations," in *Ubiquitous Intelligence and Computing*, ser. Lecture Notes in Computer Science, Z. Yu, R. Liscano, G. Chen, D. Zhang, and X. Zhou, Eds. Springer, 2010, vol. 6406, pp. 548–562.

[24] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[25] L. Breiman, "Random forests," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.

[26] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, p. 273297, 1995.

[27] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 27, pp. 1–27, 2011.