Human Activities Recognition in Android Smartphone Using WSVM-HMM Classifier

M'hamed Bilal Abidine[(✉)] and Belkacem Fergani

Laboratoire d'Ingénierie des Systèmes Intelligents et Communicants,
LISIC Lab., Electronics and Computer Sciences Department,
University of Science and Technology Houari Boumediene (USTHB),
Algiers, Algeria
abidineb@hotmail.com

Abstract. Being able to recognize human activities is essential for several applications such as health monitoring, fall detection, context-aware mobile applications. In this work, we perform the recognition of the human activity based on the combined Weighted SVM and HMM by taking advantage of the relative strengths of these two classification paradigms. One significant advantage in WSVMs is that, they deal the problem of imbalanced data but his drawback is that, they are inherently static classifiers - they do not implicitly model temporal evolution of data. HMMs have the advantage of being able to handle dynamic data with certain assumptions about stationary and independence. The experiment results on real datasets show that the proposed method possess the better robustness and distinction.

Keywords: Activity recognition · Classification · Weighted SVM · HMM

1 Introduction

The advancement of technologies has facilitated the monitoring of human activities through the embedded sensors in a smartphone. Recently, smart phones, equipped with a rich set of sensors, are explored as alternative platforms for human activity recognition (HAR) [1, 2]. HAR technology aims at recognizing the behavior and activities of users through a series of observations, which has wide application [3, 4] in different areas, such as healthcare and military monitoring.

With smartphones becoming an integral part of daily human life [5], they are being preferred as the most usable appliances that could recognize human activities due to its powerful in terms of mobility, user-friendly interface, network capability, strong CPU, memory, and battery. They contain a large number of hardware sensors such as accelerometer, gyroscope, temperature, humidity, light sensor, and GPS receiver.

The human sensor based activity recognition is a combination of sensor networks hand-in-hand with the data mining and machine learning techniques [6]. The smartphones provide enormous amount of sensor data for one to understand the daily activity patterns of an individual.

The basic procedure for mobile activity recognition involves i) collection of labelled data, i.e., associated with a specific class or activity from users that perform sample activities to be recognized ii) classification model generation by using collected

data to train and test classification algorithms iii) a model deployment stage where the learnt model is transferred to the mobile device for identifying new contiguous portions of sensor data streams that cover various activities of interest. Sensor data can be processed in real-time or logged for offline analysis and evaluation. The model generation is usually performed offline on a server system and later deployed to the phone to recognize the activity performed.

Recently, several authors [7, 8] have proposed many applications related to activity recognition on multiple body positions. Most of the work, like Ahmad [9], Tran [10], Awan [11], Shoaib [12], and Abidine [13], consider a single classifier approach to study activity recognition using smartphones. For the classification, SVMs are popular [8, 14]. It is also the case for HMMs [15] which they commonly used for time-series activity recognition. However, there is very limited number of publications in the literature that investigate the application of the WSVM classifier for smartphone data, and no one is found about applying the latter one on smartphone data or even on HAR system's datasets. Building a system with high precision to accurately identify these activities is a challenging task.

In this work, we adopted a new method for physical activity recognition using mobile phones that uses labels outputting WSVM in HMM. WSVM investigated the effect of overweighting the minority class on SVM modeling between the performed activities. HMM is a natural solution to address the activity complexity by — capturing and smoothing information during the transition between two activities (e.g. Walking and Standing). We also used the feature extraction approach that transforms the original high dimensional data to a lower dimensional feature space. The transformation can be linear or nonlinear. In this project, we employed the linear Principal Component Analysis (PCA) [16] to extract the feature vectors.

2 The Proposed HAR System by Combining WSVM-HMM Based PCA

2.1 Overview

Figure 1 shows the architecture of the proposed activity recognition system. Among the available labelled data, training and test subsets are chosen using the cross-validation mechanism. The constructed PCA space is then used for training and testing the Weighted SVM classifier. In the second step of the process is a pre-classification by 'WSVM', this phase is carried out by the 'cross-validation' will generate an estimate of the label vector.

The principal component features concatenated with the WSVM estimated label vector are employed as a new training data to train HMM classifier. The final classification is performed with the 'Viterbi' algorithm, by the use of a HMM model.

An estimated label vector is generated by the 'Viterbi' algorithm and the system will output the recognized activity (i.e., walking, running, and others).
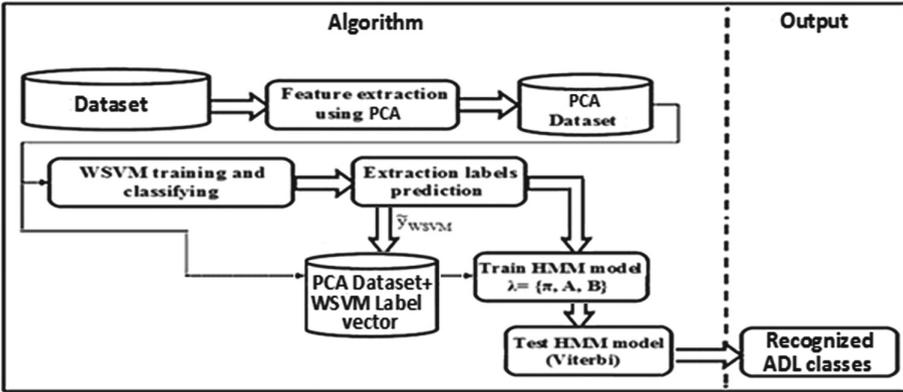
Fig. 1. Hybrid WSVM-HMM system based PCA approach.

2.2 Principal Component Analysis (PCA)

PCA [16] is an orthogonal projection-based technique such that the variance of the projected data is maximized. In our case, a large number of features are extracted by prepossessing the raw signals generated from different sensors. It is a widely used technique for dimensionality reduction, feature extraction, and data visualization through the construction of uncorrelated principal components that are a linear combination of the original variables. The PCA components can be counted by performing the eigenvector decomposition of the covariance matrix S:

$$S = \sum_{j=1}^{m} \left((\vec{x})_j - \mu \right) \left((\vec{x})_j - \mu \right)^T, \quad \mu = \frac{1}{m} \sum_{j=m}^{m} (\vec{x})_j. \tag{1}$$

This problem leads to solve the eigenvalue equation with λ *is the eigenvalue of S and V is the eigenvector corresponding to the* λ:

$$\lambda V = SV, \|V\| = 1. \tag{2}$$

Where $V = [v_1, v_2, \ldots, v_i]$, $(i = 1, \ldots, n)$ is the $n \times n$ matrix containing n eigenvectors and λ is an $n \times n$ diagonal matrix of eigenvalues of the covariance matrix. In Eq. (2), each n dimensional eigenvector v_i corresponds to the ith eigenvalue λ_i.

2.3 Weighted Support Vector Machines (WSVM)

Osuna et al. [17] proposed an extension of the SVM modeling, Weighted SVM algorithm to overcome the imbalance problem by introducing two different penalty parameter C_- and C_+ in the primal Lagrangian (Eq. 3) for the minority ($y_i = -1$) and majority classes ($y_i = +1$), as follow

$$\min_{s,b,\zeta} \frac{1}{2} w \bullet w + C_+ \sum_{i|y_i=1}^{m_+} \zeta_i + C_- \sum_{i|y_i=-1}^{m_-} \zeta_i \tag{3}$$

$$\text{subject to } y_i(s \bullet \Phi(y_i) + b) \geq 1 - \zeta_i, \ \zeta_i \geq 0, \ i = 1,\ldots,m \ .$$

m_+ *(resp. m_-)* the number of positive (resp. negative) instances in the initial database ($m_- + m_+ = m$). Solving the formulation dual of WSVM [17] gives a decision function for classifying a test point $y \in R^p$

$$f(x) = \text{sgn}\left(\sum_{i=m}^{m_{sv}} \alpha_i y_i K(x, x_i) + b \right) . \tag{4}$$

We used the Gaussian kernel as follows: $K(x,y) = \exp\left(-\|x-y\|^2/2\sigma^2\right)$. Some authors [17–19] have proposed adjusting different cost parameters to solve the imbalanced problem. To extend Weighted SVM to the multi-class scenario in order to deal with N classes (daily activities), we have shown in [20] that the cost of mis-classifying a point from the small class should be heavier than the cost for errors on the large class. They used different misclassification C_i per class, use this conclusion can get a satisfactory result. By taking $C_- = C_i$ and $C_{+ =} C$, with m_+ and m_i be the number of samples of majority classes and number of samples in the ith class, the main ratio cost value C_i for each activity can be obtained by:

$$C_i = \text{round}(C \times [m_+/m_i]), \ i = 1,\ldots,N. \tag{5}$$

2.4 Hidden Markov Model (HMM)

HMM [21] comprises two parts: Markov chain and stochastic process. Markov chain, whose output is a sequence of state, can be described by the initial probability distri-bution for the states (π) and the state transition matrix (A), while stochastic process whose output is a sequence of observed values, is described by the observation probability matrix (B). Thus, a HMM can be described as:

$$A = a_{ij} = P(y_t = j|y_{t-1} = i) \text{ and } \sum_{j=1}^{N} a_{ij} = 1 \tag{6}$$

$$B = \left[b_j(O_t) \right] \tag{7}$$

$$b_j(O_t) = P(q_{k+1} = O_t/q_t = i) \tag{8}$$

$$\pi = [\pi_1, \pi_2\ldots,\pi_N] \tag{9}$$

$$\pi_i = P(q_0 = i). \tag{10}$$

With: $i, j \in \{1,2, …, N\}$

O_t: Vector of observations

A standard HMM is a generative probabilistic model, which generates hidden states y_t from observable data x_t at each discrete time instant. In our case the hidden variable is the activities that the subject was performing at a given time step and the observable variable is the vector of sensor readings. HMM model mainly works on two basic principles as follows: the observable variable at time t, namely x_t, depends *only* on the hidden variable y_t. The hidden variable at time t, namely y_t, depends *only* on the previous hidden variable y_{t-1}.

Learning the parameters of these parameters corresponds to maximizing the joint probability p(x, y) between the sensor data and activities in the training data. The joint probability therefore factorizes as follows:

$$P(\mathrm{x}, \mathrm{y}) = \prod_{t=1}^{T} p(y_t|y_{t-1})p(x_t|y_t) . \tag{11}$$

The main aim of this model is to determine the best hidden state sequence from the observed output sequence that maximizes $p(\mathrm{x}, \mathrm{y})$.

3   Experimental Results and Analysis

3.1   Datasets

We validate our method on three public datasets whose information is summarized in Table 1. The first dataset used is from [22]: the Human Activity Dataset (HAR). The second dataset (HAPT) [23] with Postural Transitions is similar to previous dataset, further, it includes postural transitions in addition of the previous version of the dataset Records. The third dataset is from [24], titled Wireless Sensor Data Mining (WISDM). All datasets have been recorded by means of Android smartphone. For the annotation of the activities, the video-recorded is used to label the data manually. The HAR and

Table 1. Summary of datasets used in the evaluation

Houses	HAR	HAPT	WISDM
Nb of subjects	30	30	29
Annotation	Video	Video	Graphical user interface
$F_{Sampling}$ (Hz)	50	50	20
Features	561	561	6
Smartphone	Samsung Galaxy SII	Samsung Galaxy SII	Cell Phone
Position	Waist	Waist	Front leg pocket
Sensors	Accelerometer and gyroscope	Accelerometer and gyroscope	Accelerometer
Activities	6	12	6

HAPT datasets provide a large extracted features extracted by prepossessing the raw signals generated from sensors.

3.2   Results

These algorithms are tested under MATLAB environment and the WSVM algorithm is tested with implementation LibSVM [25] using Gaussian kernel is used for all the datasets. Each training dataset is normalized before classification within a range of [−1, 1]. We optimized the SVM hyper-parameters (σ, C) for all training sets in the range [0.1, 0.2, 0.5, 1] and {0.1, 1, 5, 10, 100}, respectively, to maximize the error rate of five fold- cross validation technique. The optimal parameters σ_{opt} = 0.9, 0.9, and 0.8 are found to be optimal the training dataset of HAR, HAPT, and WISDM, respectively. We show in the Table 2 that the fusion of principal component features with WSVM-HMM makes the model more robust, achieving better performance. One also notices for HAR dataset that the multi-class WSVM method improves the classification results over MC-SVM, MC-HF-SVM and HMM classifiers used alone. On the other hand, the results also show that WSVM outperforms HMM for recognizing activities for all datasets except for the HAPT dataset.

In terms of reducing the datasets, the feature reduction identifies the most relevant features for the learning process. We notice that PCA features can improve the discrimination between different activities than the original features. For WISDM the performances of activity recognition are low than HAR and HAPT datasets with 561 features. This is explained by the number of features (6) for WISDM is not sufficient when using PCA algorithm. Another reason to the lowest accuracy in WISDM dataset is attributed to the use only the accelerometer sensor comparatively to the HAR and HAPT that use the both accelerometer and gyroscope sensors.

Table 2. The micro-averaged measures: Recall, Precision, F-measure and Accuracy for all approaches in (%). Bold values are the results for our approach for each dataset.

Datasets	Approach	Recall	Precision	F-measure	Accuracy
HAR	MC-SVM [8]	89.6	89.9	89.7	89.3
	MC-HF-SVM [8]	89.3	89.2	89.2	89.0
	WSVM	92.4	91.6	91.9	93.9
	HMM	89.2	90.2	89.7	93.7
	Proposed	**94.0**	**96.7**	**95.3**	**94.9**
HAPT	WSVM	96.0	92.4	94.1	86.1
	HMM	98.3	97.1	97.7	96.5
	Proposed	**97.3**	**99.0**	**98.1**	**96.8**
WISDM	J48 [26]	81.7	–	–	85.1
	LogisticRegression [26]	68.4	–	–	78.1
	MultilayerPerceptron [26]	80.4	–	–	91.7
	WSVM	83.4	76.5	79.8	81.4
	HMM	79.4	80.0	79.7	84.9
	Proposed	**91.9**	**79.8**	**85.4**	**92.3**

Table 3. Confusion matrix of activities for the proposed method on the HAR dataset.

Activities	Walking	W. Upstairs	W. Downstairs	Sitting	Standing	Laying
Walking	**97.1**	2.1	0.7	0.0	0.0	0.1
Walking. Upstairs	1.2	**96.2**	2.4	0.1	0.1	0.0
Walking. Downstairs	0.7	2.2	**97.1**	0.0	0.0	0.0
Sitting	0.6	0.0	0.1	**83.5**	12.2	3.6
Standing	0.1	0.2	0.2	7.4	**91.4**	0.7
Laying	0.0	0.0	0.3	0.8	0.3	**98.6**

To get a detailed knowledge of the performances on each class corresponding to current activity for the HAR dataset with six different activities. We calculate the confusion matrix of the proposed method in Table 3. From these tables, we see that the best performances were obtained for the proposed method for all classes, in particular for the static activities (Sitting and Standing).

In the Table 3, 96.2% of 'W. Upstairs' activity instances are correctly recognized, while 2.4% goes into 'W. Downstairs' and 1.2% are confused with 'Walking' activity. The similar classes such as 'Walking', 'W. Upstairs', and 'W. Downstairs' show similar trend of sharing errors among each other. The reason is the similar status of smartphone when the user does these dynamic activities. We notice that the static activities share errors among each other. 12.2% of 'Standing' activity instances are confused with 'Sitting' activity and 7.4% of 'Sitting' activity instances are confused with 'Standing' activity. Intuitively, this can be explained by the fact that the patterns in the acceleration data between these activities are somewhat similar.

4   Conclusion and Future Work

Experimental results of the hybrid model presented demonstrate how it can be effectively employed for activity recognition of static and dynamic activities. It obtains a significant performance. Specifically, we show how the hybrid system obtained by using the WSVM label output a new feature added to the reduced data for training and testing HMM outperforms other well known supervised pattern recognition approaches. We consider that WSVM approach has great potential to deal the imbalance class in this human activity recognition problem. However, it must be noticed that hybridizing these schemes implies a more complex system. Fortunately, the training phase in a deployed activity recognizer is usually done offline, so we do not consider such growth of complexity a real problem in our domain.

References

1. Sarwar, M., Soomro, T.R.: Impact of smartphone's on society. Eur. J. Sci. Res. **98**(2), 216–226 (2013)
2. Abidine, M.B., Fergani, L., Fergani, B., Fleury, A.: Improving human activity recognition in smart homes. Int. J. E-Health Med. Commun. (IJEHMC) **6**(3), 19–37 (2015)

3. Klasnja, P., Pratt, W.: Healthcare in the pocket: mapping the space of mobile-phone health interventions. J. Biomed. Inf. **45**, 184–198 (2012)
4. Candás, J.L.C., Peláez, V., López, G., Fernández, M.Á., Álvarez, E., Díaz, G.: An automatic data mining method to detect abnormal human behaviour using physical activity measurements. Pervasive Mob. Comput. **15**, 228–241 (2014)
5. Shoaib, M., Bosch, S., Incel, D., Scholten, H., Havinga, P.J.M.: A survey of online activity recognition using mobile phones. J. Sens. (Basel, Switzerland) **15**(1), 2059–2085 (2015)
6. Fu, Y.: Human Activity Recognition and Prediction. Springer, Switzerland (2016). https://doi.org/10.1007/978-3-319-27004-3
7. Chetty, G., White, M., Akther, F.: Smart phone based data mining for human activity recognition, Elsevier Procedia Comput. Sci. – ICICT **46**(8), 1181–1187 (2015)
8. Anguita, D., Ghio, A., Oneto, L., Parra, X., Reyes-Ortiz, J.L.: Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In: Bravo, J., Hervás, R., Rodríguez, M. (eds.) IWAAL 2012. LNCS, vol. 7657, pp. 216–223. Springer, Heidelberg (2012). https://doi.org/10.1007/978-3-642-35395-6_30
9. Ahmad, N., Han, L., Iqbal, K., Ahmad, R., Abid, M.A., Iqbal, N.: SARM: salah activities recognition model based on smartphone. Electronics **8**(8), 881 (2019)
10. Tran, D.N., Phan, D.D.: Human activities recognition in android smartphone using support vector machine. In: 2016 7th International Conference on Intelligent Systems, Modelling and Simulation (ISMS), pp. 64–68. IEEE (2016)
11. Awan, M.A., Guangbin, Z., Kim, H.C., Kim, S.D.: Subject-independent human activity recognition using smartphone accelerometer with cloud support. Int. J. Ad Hoc Ubiq. Comput. **20**(3), 172–185 (2015)
12. Shoaib, M., Scholten, H., Havinga, P.J.M.: Towards physical activity recognition using smartphone sensors. In: Proceedings of the 2013 IEEE 10th International Conference on Ubiquitous Intelligence and Computing and IEEE 10th International Conference on Autonomic and Trusted Computing, Italy 18–21, pp. 80–87 (2013)
13. Abidine, M.B., Yala, N., Fergani, B., Clavier, L.: Soft margin SVM modeling for handling imbalanced human activity datasets in multiple homes. In: 2014 International Conference on Multimedia Computing and Systems (ICMCS), pp. 421–426. IEEE (2014)
14. Alman, A., Lawi, A., Tahir, Z.: Pattern recognition of human activity based on smartphone data sensors using SVM multiclass. In: 1st International Conference on Science and Technology, ICOST 2019. European Alliance for Innovation (EAI) (2019)
15. Cheng, B.C., Tsai, Y.A., Liao, G.T., Byeon, E.S.: HMM machine learning and inference for Activities of Daily Living recognition. J. Supercomput. **54**(1), 29–42 (2010)
16. Jolliffe, I.T.: Principal Component Analysis, 2nd edn. Springer, NewYork (2010)
17. Osuna, E., Freund, R., Girosi, F.: Support vector machines: training and applications. Massachusetts Institute of Technology, Cambridge (1997)
18. Veropoulos, K., Campbell, C., Cristianini, N.: Controlling the sensitivity of support vector machines. In: Proceedings of the International Joint Conference on AI, pp. 55–60 (1999)
19. Huang, Y.M., Du, S.X.: Weighted support vector machine for classification with uneven training class sizes. In: Proceedings of the IEEE International Conference on Machine Learning and Cybernetics, vol. 7, pp. 4365–4369 (2005)
20. Abidine, B.M., Fergani, L., Fergani, B., Oussalah, M.: The joint use of sequence features combination and modified weighted SVM for improving daily activity recognition. Pattern Anal. Appl. **21**(1), 119–138 (2016). https://doi.org/10.1007/s10044-016-0570-y
21. Cheng, B.C., Tsai, Y.A., Liao, G.T., Byeon, E.S.: HMM machine learning and inference for activities of daily living recognition. J. Supercomput. **54**(1), 29–42 (2010). https://doi.org/10.1007/s11227-009-0335-0
22. https://archive.ics.uci.edu/ml/machine-learning-databases/00240/. Accessed Mar 2017

23. https://archive.ics.uci.edu/ml/datasets/SmartphoneBased+Recognition+of+Human +Activities+and+Postural+Transitions Accessed 10 Mar 2016
24. http://www.cis.fordham.edu/wisdm/dataset.php. Accessed Mar 2017
25. Hsu, C.W., Chang, C.C., Lin, C.J.: A practical guide to support vector classification (2003). In: Chang, C.-C., Lin, C.-J.: 'LIBSVM: A library for support vector machines' ACM Transactions on Intelligent Systems and Technology, vol. 2, pp. 27:1–27:27 (2011). http://www.csie.ntuedu.tw/cjlin/libsvm
26. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. ACM SIGKDD Explor. Newslett. **12**(2), 74–82 (2011)