Training Computationally Efficient Smartphone–Based Human Activity Recognition Models*

Davide Anguita[1], Alessandro Ghio[1], Luca Oneto[1], Xavier Parra[2], and Jorge Luis Reyes-Ortiz[1,2]

[1] DITEN – University of Genoa, Via Opera Pia 11A, Genova, I-16145, Italy
{Davide.Anguita,Alessandro.Ghio,Luca.Oneto}@unige.it
[2] CETpD - Universitat Politècnica de Catalunya, Vilanova i la Geltrú 08800, Spain
{xavier.parra@upc.edu,jorge.luis.reyes@estudiant.upc.edu}

Abstract. The exploitation of smartphones for Human Activity Recognition (HAR) has been an active research area in which the development of fast and efficient Machine Learning approaches is crucial for preserving battery life and reducing computational requirements. In this work, we present a HAR system which incorporates smartphone-embedded inertial sensors and uses Support Vector Machines (SVM) for the classification of Activities of Daily Living (ADL). By exploiting a publicly available benchmark HAR dataset, we show the benefits of adding smartphones gyroscope signals into the recognition system against the traditional accelerometer-based approach, and explore two feature selection mechanisms for allowing a radically faster recognition: the utilization of exclusively time domain features and the adaptation of the L1 SVM model which performs comparably to non-linear approaches while neglecting a large number of non-informative features.

Keywords: Smartphones, Human Activity Recognition, SVM, L1 SVM, Feature Selection.

1 Introduction

Human Activity Recognition using wearable systems is attracting a growing interest due to the recent advances in technology, which are making available a wide range of microelectromechanical sensors and portable computing devices. It has mainly addressed healthcare-related applications such as the development of automated assisted living and ambulatory monitoring technologies for people with disabilities, going through rehabilitation, or the elderly, partially substituting the continuous supervision needed from caretakers or other family members [25,16]. HAR systems facilitate to better understand different aspects of the individuals daily living to provide context information suitable for

* This work was supported in part by the Erasmus Mundus Joint Doctorate in Interactive and Cognitive Environments, which is funded by the EACEA Agency of the European Commission under EMJD ICE FPA n 2010-0012.

V. Mladenov et al. (Eds.): ICANN 2013, LNCS 8131, pp. 426–433, 2013.

a range of applications and services. This is commonly done through the use of body sensors that measure various attributes derived from motion, location, physiological signals and the environment (e.g. accelerometers, GPS, hear rate monitors, microphones) and then by interpreting this sensory input data with Machine Learning approaches such as Artificial Neural Networks (ANN), SVMs, k-Nearest Neighbors (KNN) and the Decision Tree algorithm C4.5 [15,17].

Some recent HAR approaches have taken advantage of smartphones due to their powerful processing capabilities and opportunistic sensing [23] from the available embedded components. These smart devices have been massively produced integrating small scale and low cost inertial sensors aiming to enhance human-computer interaction (e.g. for gaming and user interfaces). Research on HAR exploiting smartphone motion sensors has mostly incorporated only accelerometers [12,15], as they have been earlier introduced in 2007 [11]. Gyroscopes have been more recently included in 2010, though some less recent approaches have considered special purpose devices for HAR purposes (e.g. [2,13]). Recently, in [29] a hybrid accelerometer and gyroscope approach was used for the classification of 9 activities using an iPhone 4. They showed insights of the benefits of adding gyroscope signals into the recognition system achieving improvements ranging from 3.1% to 13.4% in classification accuracy, though a limited set of features from the gyroscope signals (signal mean, standard deviation and sum of signal magnitude in a sliding window) was used. A major limitation for the smartphone implementation was due to the exploitation of a KNN classifier, which could become inadequate in such applications due to its demanding computations for prediction, particularly with large training sets.

Open rooms for improving smartphone-based HAR consequently exist: (i) the benefits of introducing a larger set of gyroscope-based signals should be more thoroughly evaluated; on the other hand, (ii) a proper selection of the most useful features and simple (though effective) models should be carried out, so to make HAR more suitable for devices with limited battery life and computational restrictions. This paper targets both these two issues. Regarding point (i), we exploit the HAR dataset [3], which contains gyroscope measures plus a large set of previously suggested features in the time and frequency domain [5,14,1,9,8], for our analyses. Concerning issue (ii), we resort to effective Support Vector Machine (SVM) classifiers [28] and we implement two feature selection mechanisms to allow faster and computationally non-intensive recognition: on the one hand, conversely to [18], we do not discriminate on proposed feature sets, but instead on sensor type and domain; on the other hand, L1 SVM models [27] will be implemented, allowing to perform an automatic selection of significant features emerging from the training set while keeping the appealing classification performance of conventional SVMs. However, one of the main drawbacks of L1 SVMs consists in the impossibility of resorting to non-linearity to improve classification accuracy through the kernel trick [28]: for targeting this issue, we compare the use of conventional non-linear models with the one of linear classifiers in the particular case of HAR, showing how the latter ones are characterized by better performance/complexity ratios than the former ones.

2 HAR Dataset

The Human Activity Recognition Dataset [3], that will be used in the forthcoming analyses and is available at [6], was developed by the authors for the experimentation with smartphone inertial sensors and the classification of Activities of Daily Living (ADL): *standing, sitting, laying, walking, walking upstairs* and *walking downstairs*, that we respectively define A1-A6 for the sake of simplicity.

To create the dataset, a group of 30 volunteers followed a protocol of activities while carrying an Android OS smartphone attached to a belt on their waist. The dataset includes 10299 patterns which have been divided in training and test sets in a proportion of 70% to 30%. Each pattern is represented with a feature vector of 561 elements composed of time and the frequency domain features extracted from the accelerometer and gyroscope signals. These were first preprocessed for noise reduction and the removal of the gravitational component from the acceleration signal. Then fixed-width sliding windows of 2.56 sec width and 50% overlap between them were extracted from the inertial signals. From each window, a vector of features was calculated including features such as mean, standard deviation, signal magnitude area, interquartile range, auto-regresion coefficients, largest Fast Fourier Transform (FFT) power spectrum component and correlation coefficients between signal pairs, etc. For further details, the reader can also refer to [3].

3 An Effective Smartphone-Based Solution for HAR

Our target is to design a model, which can be effectively run on smartphones with limited battery life and computational restrictions. We have thus to identify the simplest possible classifier exploiting the smallest set of features that guarantees the best performance/computational burden ratio. For these purposes, we peruse the exploitation of linear models, which use only those selected inputs that are crucial to attain sufficient classification accuracy. The following subsections are devoted to describing the analyses, that have been performed on the previously introduced dataset.

3.1 Non-linear vs. Linear SVMs

SVM is one of the most widely employed Machine Learning algorithms [28,26,7]. It allows to solve binary problems by deterministically finding a maximum-margin hyperplane that separates the data. The effectiveness of SVM models make them particularly appealing in several and heterogeneous real-world problems, including applications on smartphones [4].

Table 1. Confusion Matrix for the comparison of performance of linear and Gaussian L2 SVM

	Linear SVM							**Gaussian SVM**						
	A1	A2	A3	A4	A5	A6	%	A1	A2	A3	A4	A5	A6	%
A1	**492**	1	3	0	0	0	*99.2*	**492**	1	3	0	0	0	*99.2*
A2	18	**451**	2	0	0	0	*95.7*	18	**451**	2	0	0	0	*95.7*
A3	4	6	**410**	0	0	0	*97.6*	4	6	**410**	0	0	0	*97.6*
A4	0	2	0	**434**	55	0	*88.4*	0	2	0	**432**	57	0	*88.0*
A5	0	0	0	14	**518**	0	*97.4*	0	0	0	14	**518**	0	*97.4*
A6	0	0	0	0	0	**537**	*100.0*	0	0	0	0	0	**537**	*100.0*
%	*95.7*	*98.0*	*98.8*	*96.9*	*90.4*	*100.0*	***96.4***	*95.7*	*98.0*	*98.8*	*96.9*	*90.1*	*100.0*	***96.4***

Given a dataset $\{(\boldsymbol{x}_1, y_1), \ldots, (\boldsymbol{x}_l, y_l)\}$, $\boldsymbol{x}_i \in \mathbb{R}^m$, and $y_i \in \{\pm 1\}$, an SVM classifier $f(\boldsymbol{x}) = \boldsymbol{w} \cdot \boldsymbol{x} + b$ is found by solving the following primal problem, which exploits an L2 regularization term to adjust the size of the class of functions [28]:

$$\begin{aligned} \min_{\boldsymbol{w}, b, \boldsymbol{\xi}} \quad & \frac{1}{2} \|\boldsymbol{w}\|_2^2 + C \mathbf{1}_l^T \boldsymbol{\xi} \\ \text{s.t.} \quad & Y (X\boldsymbol{w} + \boldsymbol{b}_l) \geq \mathbf{1}_l - \boldsymbol{\xi}, \; \boldsymbol{\xi} \geq \mathbf{0}_l, \end{aligned} \tag{1}$$

where $\xi_i = \max[0, (1 - y_i f(\boldsymbol{x}_i))]$, $X = [\boldsymbol{x}_1 | \ldots | \boldsymbol{x}_l]^T$, $\boldsymbol{y} = [y_1 | \ldots | y_l]^T$, $Y = \text{diag}(\boldsymbol{y})$ (Y is a diagonal matrix where the element on the diagonal are the $y_{i \in \{1, \ldots, n\}}$), and $\boldsymbol{a}_p$ is a vector of p elements all equal to a. By introducing n Lagrange multipliers $\boldsymbol{\alpha}$ we obtain the dual formulation of the conventional SVM:

$$\begin{aligned} \min_{\boldsymbol{\alpha}} \quad & \frac{1}{2} \boldsymbol{\alpha}^T Q \boldsymbol{\alpha} - \mathbf{1}_l^T \boldsymbol{\alpha} \\ \text{s.t.} \quad & \mathbf{0}_l \leq \boldsymbol{\alpha} \leq \boldsymbol{C}_l, \boldsymbol{y}^T \boldsymbol{\alpha} = 0, \end{aligned} \tag{2}$$

where $Q \in \mathbb{R}^{l \times l}$ and $Q_{ij} = y_i y_j K(\boldsymbol{x}_i, \boldsymbol{x}_j)$, where $K(\cdot, \cdot)$ is the kernel function. As we are targeting multiclass classification problems, generalization through the One-Vs.-All (OVA) approach is exploited [22,20].

The first performed experiments aim at comparing the performance of SVM models, based on linear and Gaussian kernels. The confusion matrices in Table 1 depict the classification results given the 6 ADL using the complete set of features introduced in Section 2. The accuracies achieved with the two methods are substantially identical, thus showing the equivalence between these two models. The linear approach is consequently preferred for the prediction of activities, more specifically for its application in limited resources devices: in fact, the prediction phase is much faster than the kernelized approach and linear models allow to exploit more sophisticated dimensionality reduction approaches, as will be shown in the next sections. Results also show balanced precision and recall measures for all the activities.

Table 2. Experiments with different feature subsets and conventional linear SVM models

Acc	Gyro	Time	Freq	Feature groups	N. Features	Linear SVM
0	1	1	0	GT	108	78.0%
0	1	1	1	GTF	213	81.0%
1	0	1	0	AT	164	90.6%
1	0	1	1	ATF	348	91.2%
1	1	1	0	AGT	272	95.8%
1	1	1	1	AGTF	561	96.4%

3.2 Selection of Subsets of Features

The second set of analyses consists in evaluating the inputs available in the dataset aiming at a reduction in the number of significant features. This is achieved by separating the inputs in groups with respect to: the type of sensor employed, namely *Accelerometer* (A) and *Gyroscope* (G); the domain, namely *Time* (T) and *Frequency* (F).

In practice, we expect to balance the trade-off between the addition of meaningful features and the removal of the ones that are redundant or that require expensive computations for their estimation. For such purposes, we test all the combinations of feature groups and compute the test error rate performed by a linear SVM model, trained on the subset employed: Table 2 presents the results. They suggest that the whole set of features (AGTF) should be exploited, although frequency-related inputs are not strictly necessary for this application as they do not largely affect recognition performance while, contrarily, requiring a remarkable computational effort for their derivation. Results also allow to gather some evidence of the benefits, which the addition of gyroscope signals bring into the HAR system, thus counterbalancing the (limited) slowdown in prediction due to the presence of these extra features. Note however that the models trained with sets using only gyroscope features (GT, GTF) have a lower performance, suggesting that the use of this sensor by its own is not appropriate for this application, despite enhancing the recognition when exploited concurrently with accelerometers.

3.3 Dimensionality Reduction with L1-SVM

As the conventional SVM does not perform any dimensionality reduction [10], which is desirable in some practical applications to highlight relevant features as well as to reduce the computational burden of performing the classification of new samples, the replacement of the L2 term with an L1 Manhattan norm based regularization has been proposed (L1 SVM) [27]:

$$\begin{aligned} \min_{\boldsymbol{w},b,\boldsymbol{\xi}} \quad & \|\boldsymbol{w}\|_1 + C\mathbf{1}_l^T\boldsymbol{\xi}, \\ \text{s.t.} \quad & Y\left(X\boldsymbol{w} + \boldsymbol{b}_l\right) \geq \mathbf{1}_l - \boldsymbol{\xi},\ \ \boldsymbol{\xi} \geq \mathbf{0}_l. \end{aligned} \tag{3}$$

Table 3. Comparison of results using L2 and L1 SVM models

Feature groups	N. Features	L2 (linear) SVM	L1 SVM	N. Features L1 SVM	ρ
AGT	272	95.8%	96.5%	174	64.0 %
AGTF	561	96.4%	96.5%	275	49.0 %

Note that the conventional kernel trick cannot be exploited in the previous formulation, thus the effectiveness of linear models assumes an ever greater importance. This problem is a standard Linear Programming problem, for which many tools have been developed throughout the years [19].

L1 SVM allows to perform dimensionality reduction thanks to the characteristics of the Manhattan norm exploited, that is several weights w_i will be (generally) nullified during the learning procedure: this is in contrast with the conventional (L2) SVM, where $w_i \neq 0 \; \forall i = 1, ..., m$ in the final model. As we are exploiting OVA for targeting multiclass classification, we will consider a feature as filtered if the corresponding weight is null for all the (six) models learned for the OVA approach.

Table 3 presents the comparison of linear (L2) SVMs and L1 SVMs, both in terms of accuracy and number of selected features (remembering that L2 procedures do not perform any dimensionality reduction). In particular, we considered only the groups of features that showed to be necessary for HAR purposes, according to the results derived in the previous section. It is worth noting that L1 models perform comparably to (and, unexpectedly from literature, slightly better than) L2 models, furthermore allowing to remarkably reduce the dimensionality of the problem. The remarkable classification performance of L1 models is probably due to the filtering of noisy features, which negatively afflict L2 classifiers.

As a final remark, the results obtained with the L1 SVM algorithm also outperform the ones obtained at the ESANN 2013 HAR competition [21] in which contestants were challenged to propose novel approaches for the recognition of activities using the same HAR dataset. Linear and non linear Machine Learning methods were proposed, achieving a maximum classification accuracy of 96.4% with the work presented in [24], where an One-Vs.-One (OVO) SVM classification approach [22] was employed for the recognition task.

4 Conclusions

In this paper, we showed the benefits of adding gyroscope information into a human activity recognition system based on smartphone technology. We verify that a set of common daily activities can be accurately classified when this sensor is used along with the accelerometer. We explored three SVM algorithms including linear (L1 SVM, conventional L2 linear SVM) and non-linear (L2 Gaussian SVM) approaches on the HAR dataset and found a similar performance in terms of accuracy, but our selection criterion was subject to prediction speed and the possibility of applying them in devices with limited resources to provide less computational complexity and energy consumption.

Linear approaches exhibited the best trade off between accuracy and prediction speed, conferring distinctive benefits to the L1 SVM, which provides itself a reduction of the effective number of features needed for the prediction of the ADL. Furthermore, the study between different feature domains lead us to disregard frequency domain features as they were not only marginally contributing to the recognition performance but also adding expensive computations for their estimation. The ideal model selected for our application was the AGT, which only takes into account time domain features from the accelerometer and the gyroscope.

Future work will explore novel model selection approaches on the L1 SVM algorithm, that can help to further reduce the number of effective features by considering near-optimal hyperparameter models within the OVA binary classifiers to increase the number of zero-valued weights matches.

References

1. Allen, F.R., Ambikairajah, E., Lovell, N.H., Celler, B.G.: Classification of a known sequence of motions and postures from accelerometry data using adapted gaussian mixture models. Physiological Measurement 27(10), 901–935 (2006)
2. Altun, K., Barshan, B.: Human activity recognition using inertial/Magnetic sensor units. In: Salah, A.A., Gevers, T., Sebe, N., Vinciarelli, A. (eds.) HBU 2010. LNCS, vol. 6219, pp. 38–51. Springer, Heidelberg (2010)
3. Anguita, D., Ghio, A., Oneto, L., Parra, X., Reyes-Ortiz, J.L.: A public domain dataset for human activity recognition using smartphones. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN (2013)
4. 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)
5. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
6. Frank, A., Asuncion, A.: UCI machine learning repository (2010)
7. Ghio, A., Anguita, D., Oneto, L., Ridella, S., Schatten, C.: Nested sequential minimal optimization for support vector machines. In: Villa, A.E.P., Duch, W., Érdi, P., Masulli, F., Palm, G. (eds.) ICANN 2012, Part II. LNCS, vol. 7553, pp. 156–163. Springer, Heidelberg (2012)
8. Karantonis, D.M., Narayanan, M.R., Mathie, M., Lovell, N.H., Celler, B.G.: Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Transactions on Information Technology in Biomedicine 10(1), 156–167 (2006)
9. Khan, A.M., Lee, Y.K., Lee, S., Kim, T.S.: Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In: IEEE International Conference on Future Information Technology (2010)
10. Khan, N.M., Ksantini, R., Ahmad, I.S., Guan, L.: A sparse support vector machine classifier with nonparametric discriminants. In: Villa, A.E.P., Duch, W., Érdi, P., Masulli, F., Palm, G. (eds.) ICANN 2012, Part II. LNCS, vol. 7553, pp. 330–338. Springer, Heidelberg (2012)

11. Lane, N.D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., Campbell, A.T.: A survey of mobile phone sensing. IEEE Communications Magazine 48(9), 140–150 (2010)
12. Lara, O., Labrador, M.: A survey on human activity recognition using wearable sensors. IEEE Communications Surveys Tutorials 1(99), 1–18 (2012)
13. Lee, S.W., Mase, K.: Activity and location recognition using wearable sensors. IEEE Pervasive Computing 1(3), 24–32 (2002)
14. Lovell, N., Wang, N., Ambikairajah, E., Celler, B.G.: Accelerometry based classification of walking patterns using time-frequency analysis. In: IEEE Annual International Conference of the Engineering in Medicine and Biology Society (2007)
15. Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors 10(2), 1154–1175 (2010)
16. Narayanan, M.R., Scalzi, M.E., Redmond, S.J., Lord, S.R., Celler, B.G., Lovell, N.H.: A wearable triaxial accelerometry system for longitudinal assessment of falls risk. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society (2008)
17. Nishkam, R., Nikhil, D., Preetham, M., Littman, M.L.: Activity recognition from accelerometer data. In: Conference on Innovative Applications of Artificial Intelligence (2005)
18. Preece, S.J., Goulermas, J.Y., Kenney, L.P., Howard, D.: A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. IEEE Transactions on Biomedical Engineering 56(3), 871–879 (2009)
19. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P.: Numerical recipes: The art of scientific computing, 3rd edn. Cambridge University Press (2007)
20. Ramírez, F., Allende, H.: Dual support vector domain description for imbalanced classification. In: Villa, A.E.P., Duch, W., Érdi, P., Masulli, F., Palm, G. (eds.) ICANN 2012, Part I. LNCS, vol. 7552, pp. 710–717. Springer, Heidelberg (2012)
21. Reyes-Ortiz, J.L., Ghio, A., Anguita, D., Parra, X., Cabestany, J., Catal, A.: Human activity and motion disorder recognition: Towards smarter interactive cognitive environments. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN (2013)
22. Rifkin, R., Klautau, A.: In defense of one-vs-all classification. The Journal of Machine Learning Research 5, 101–141 (2004)
23. Roggen, D., Förster, K., Calatroni, A., Holleczek, T., Fang, Y., Tröster, G., Lukowicz, P., Pirkl, G., Bannach, D., Kunze, K., Ferscha, A., Holzmann, C., Riener, A., Chavarriaga, R., del, R., Millán, J.: Opportunity: Towards opportunistic activity and context recognition systems. In: Proc. 3rd IEEE WoWMoM Workshop on Autononomic and Opportunistic Communications (2009)
24. Romera-Paredes, B., Aung, H., Bianchi-Berthouze, N.: A one-vs-one classifier ensemble with majority voting for activity recognition. In: European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN (2013)
25. Ryder, J., Longstaff, B., Reddy, S., Estrin, D.: Ambulation: A tool for monitoring mobility patterns over time using mobile phones. In: IEEE International Conference on Computational Science and Engineering (2009)
26. Schölkopf, B., Smola, A.J.: Learning with kernels: Support vector machines, regularization, optimization, and beyond. MIT press (2001)
27. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological), 267–288 (1996)
28. Vapnik, V.: Statistical learning theory. Wiley-Interscience (1998)
29. Wu, W., Dasgupta, S., Ramirez, E.E., Peterson, C., Norman, G.J.: Classification accuracies of physical activities using smartphone motion sensors. Journal of Medical Internet Research 14(5), 105–130 (2012)