ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/306519729

Activity recognition based on inertial sensors for Ambient Assisted Living

Conference Paper · July 2016

CITATIONS
24

READS
287

7 authors, including:

Kadian Davis-Owusu
Delft University of Technology
12 PUBLICATIONS **67** CITATIONS
SEE PROFILE

Evans Owusu
8 PUBLICATIONS **58** CITATIONS
SEE PROFILE

Vahid Bastani
Università degli Studi di Genova
13 PUBLICATIONS **125** CITATIONS
SEE PROFILE

Lucio Marcenaro
Università degli Studi di Genova
172 PUBLICATIONS **1,294** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project BioSoundscape: Nature Soundscape Based Auditory Biofeedback for Stress Mitigation View project

Project Enrich Biofeedback Interaction View project

All content following this page was uploaded by Kadian Davis-Owusu on 12 February 2017.

The user has requested enhancement of the downloaded file.

19th International Conference on Information Fusion
Heidelberg, Germany - July 5-8, 2016

Activity Recognition Based on Inertial Sensors for Ambient Assisted Living

Kadian Davis*, Evans Owusu‡, Vahid Bastani†, Lucio Marcenaro †, Jun Hu *, Carlo Regazzoni † and Loe Feijs*

*Department of Industrial Design, Eindhoven University of Technology, Eindhoven, Netherlands
Email: {K.A.Davis, J.Hu, L.M.G.Feijs}@tue.nl

†Department of Electrical, Electronic, Telecommunications Engineering and Naval Architecture (DITEN), University of Genova, Genoa - Italy
Email: {vahid.bastani, sulman.baig}@ginevra.dibe.unige.it, {carlo.regazzoni, lucio.marcenaro}@unige.it

‡Independent Researcher, Eindhoven Netherlands
Email: owboateng@gmail.com

Abstract—**Ambient Assisted Living (AAL) aims to create innovative technical solutions and services to support independent living among older adults, improve their quality of life and reduce the costs associated with health and social care. AAL systems provide health monitoring through sensor based technologies to preserve health and functional ability and facilitate social support for the ageing population. Human activity recognition (HAR) is an enabler for the development of robust AAL solutions, especially in safety critical environments. Therefore, HAR models applied within this domain (e.g. for fall detection or for providing contextual information to caregivers) need to be accurate to assist in developing reliable support systems. In this paper, we evaluate three machine learning algorithms, namely Support Vector Machine (SVM), a hybrid of Hidden Markov Models (HMM) and SVM (SVM-HMM) and Artificial Neural Networks (ANNs) applied on a dataset collected between the elderly and their caregiver counterparts. Detected activities will later serve as inputs to a bidirectional activity awareness system for increasing social connectedness. Results show high classification performances for all three algorithms. Specifically, the SVM-HMM hybrid demonstrates the best classification performance. In addition to this, we make our dataset publicly available for use by the machine learning community.**

I. Introduction

Life expectancy is increasing, which is primarily influenced by medical advances in the diagnosis and treatment of diseases, a causative factor for mankind's massive survival to old age. It is estimated that over 2 billion people will be over 60 years by 2050 [1]. As persons age, they become vulnerable to motor and physical disabilities, which limit their participation in common activities of daily living [2]. Furthermore, socially and physically inactive seniors are more susceptible to chronic diseases and social disconnectedness, causing negative outcomes on mental and cardiovascular health and mortality [3], [4]. This increases expenditures on long-term care; thus health and social care services are searching for alternatives to cope with the macro-economic challenges associated with an increasing ageing population.

In recent times, human activity recognition (HAR) has been a key component of ambient assisted living (AAL) applications for recognizing activities of daily living (ADL) [5], fall detection [6] and monitoring physical activity levels [7] for sustaining quality of life and independent living among older people. Unlike previous studies, which have been largely focused on elderly ambulatory monitoring for emergency detection, our ultimate research goal is centered around improving interpersonal awareness and social connectedness i.e. the sense of belonging based on the feeling of self-assuredness knowing that one has enough social contacts [8], through subtle awareness. To do this, we will exploit the pre-attentive features discussed in [9], (namely colour, form, spatial position and motion) to provide a richer and better abstraction of the data. Our bidirectional ambient display system opportunistically exploits motion data from the smartphone's accelerometer and gyroscope sensors to infer six basic activities (walking, ascending the stairs and descending the stairs, sitting, standing and laying) of both user groups in their natural environment. These six activity classes are commonly exploited in human activity recognition experiments [10], [11].

A plethora of studies have proposed various methods to address the activity monitoring problem, ranging from video cameras [12], [13], wearable sensors [14] and wireless sensor networks [15]. Smartphones are considered promising solutions for enabling human activity recognition and health monitoring in AAL due to its portability, inertial sensors (accelerometer and gyroscope), communication features (WIFI, 3G and Bluetooth) and low cost [16–18].

In this paper, we investigate three activity recognition approaches, which forms part of a larger experiment [18], [19], in which an elderly and their caregiver can perceive each others activity states through a subtle ambient intelligent application to improve social connectedness. The best performing activity recognition model will be used in this bidirectional subtle awareness system. Our envisioned outcome is to demonstrate enhanced benefits for the abstract presentation of activity information using an ambient display to support social connectedness and enable peace of mind as demonstrated in [20], [21]. The application of human activity recognition (HAR) in the ambient assisted living domain is motivated by the assumption that subtle awareness of activity states could elicit affective responses and provide valuable information about health, moods and habits. In general, the safety of older adults living alone is a critical requirement of AAL applications to avoid system failure [22]. As such, algorithms deployed in this domain should maintain high reliability.

With the motivation of achieving a higher classification accuracy among all classes, we exploited the Support Vector Machines model described in [23] with the use of smartphone sensors (accelerometer and gyroscope). A hybrid of the Support Vector Machine (SVM) and Hidden Markov Model (HMM) algorithms and Artificial Neural Networks (ANNs) were also explored. Experimental results on real-life data show classification performances of over 90% for all three approaches, with SVM-HMM achieving the highest detection accuracy. This offers a lot of potential for our awareness system for improving context awareness and social connectedness for older adults and their caregivers over mediated environments.

This paper is organized as follows: a summary of the state of the art is presented in section 2. Section 3 describes the procedure and the proposed methodologies. The results are discussed in section 4. In section 5, we present the insights gained and our future research directions.

II. Related Work

The smartphone's accelerometer and gyroscope sensors have been used extensively to support human activity recognition for enabling context awareness [10], [24]. Signals recorded by mobile accelerometers and gyroscopes are typically represented in the form of time-series i.e. a sequence of data points usually collected at regular intervals [25]. Specifically, common activities such as walking, standing and laying are generally represented by time-series patterns, useful for assessing physical and cognitive well-being in ambient assisted living (AAL) environments.

In the AAL domain, different machine learning algorithms have been successfully deployed for recognizing human activities in various context including remote monitoring, detecting anomalies and promoting health and well-being. Moreover, the most popular approaches, which have been considered include supervised learning techniques [26], conditional random field [27], rule-based reasoning [28], artificial neural networks [29] and probabilistic modelling [30], [31]. Also, unsupervised learning methods have been evaluated for activity recognition [32]. Despite the success of the aforementioned techniques, sensor-based activity recognition remains challenging due to several factors including its inherently noisy nature, people performing activities differently and in different sequences and the ambiguity of sensor data. Notwithstanding these challenges, probabilistic activity recognition models have been reported to successfully handle these uncertainties [33]. In particular, Hidden Markov Models (HMM) have demonstrated solid potential for addressing the ambiguities of interpretation within the AAL domain [30].

Hidden Markov Models (HMM) are predominantly useful for activity recognition due to their capability of exploiting the temporal and sequential characteristics of activity data; thus enabling the prediction of future states from current observation data. Although HMMs have demonstrated remarkable success, they are not without limitations [34]. First, it has difficulty representing concurrent or interleaved activities, which can be problematic when modelling continuous activities within the AAL domain. Second, the HMM's strict independence assumptions make it inadequate for capturing transitive dependencies of the observations. Furthermore, it is difficult to model the feature vector extracted from accelerometer and gyroscope sensors, making it unfeasible.

A plausible solution is to use a discriminative model such as the SVM to determine the emission probabilities of an HMM, which can be combined with the dynamic temporal features of the HMM to offer improved classification accuracy in dynamic pattern recognition tasks for AAL applications as successfully deployed in [35]. In particular, the authors in [35] demonstrated that the hybrid SVM-HMM model achieves better performance when contrasted with stand alone SVM and ANN classifiers. However, this success in recognition accuracy was achieved using a network of binary sensors, which is different from the goal of using inbuilt smartphone sensors in this work. In addition, by using the HMM-SVM technique the authors in [36] recorded an overall accuracy of 96% for activity recognition using wearable devices. However, the method used for data collection was somewhat obtrusive as participants wore many wearable sensors on various parts of the body including the thigh, both wrists and neck. For user comfort, in this work, we deploy as little sensors as possible for activity recognition. Moreover, hybrid SVM-HMM approaches have been successfully applied within other domains such as speech recognition [37], speech emotion recognition [38] and analysis of facial expressions [39].

The biologically inspired Artificial Neural Networks (ANNs) are commonly presented as a collection of interconnected neurons grouped in layers, which are capable of automatic learning based on experience and approximating a non-linear combinations of features for pattern recognition [40]. Artificial Neural Networks are shown to perform well in [40], [41] for learning static (e.g. standing) and dynamic activities (e.g. walking) using a wrist-worn wireless sensing triaxial accelerometer. From this, we see that ANNs provide an efficient, robust and well-suited design methodology for pattern recognition and classification involving uncertain and complex data. However, they have some limitations, including the requirement of a large volume of training data and the difficulty of deriving an explicit model as the underlying reason for high recognition validity is often unknown [40].

In sum, many studies have been proposed in recent years to recognize physical activities based on smartphone accelerometer data using a combination of different reasoning techniques. In this paper, we developed and explored three human activity models with the aim of choosing the most accurate model for our social connectedness experiment in AAL environments. To the best of our knowledge, an activity recognition system for improving social connectedness using the smartphone's inertial sensors is missing within the AAL domain.

III. Experimental Protocol

The Samsung Galaxy S II smartphones, with inbuilt accelerometer and gyroscope sensors, were used to conduct our experiment as proposed in [18]. Our mobile sensing application was developed using Android Development tools. Signal pre-processing, feature extraction, feature selection and classification were implemented using Matlab.

A. Data Collection and Feature Extraction

In addition to publicly available smartphone activity datasets, we collected our own datasets in order to reduce the uncertainties of the former. We received 5744 samples from 31 healthy volunteers, ranging from 22 to 79 years from 14 countries, namely Russia, Italy, The Netherlands, Germany, Iran, China, India, Pakistan, Nigeria, Ghana, Tunisia, Lebanon, Jamaica and Colombia. Like it was done in [42], users were asked to perform six basic activities (walking, walking up and downstairs, standing, sitting and laying) while wearing a waist-mounted smartphone belt on their left or right side. Each activity was performed for one minute in the context of the elderlys' homes and the caregivers' working environment simulating a semi-naturalistic environment. In addition, accelerometric and gyroscopic data were collected at a sampling rate of 50Hz.

To reduce the biases associated with using our own dataset, we merged our collected dataset with the public dataset for HAR using smartphones [42], which was collected in a similar manner and at the same frequency. In total, 16043 samples were available for training, cross-validation and testing. Our dataset is publicly available on Github [43] and has also been submitted to UCI Machine Learning repository.

A method similar to [42] has been employed for extracting features from the accelerometer and gyroscope data collected. Features were computed on a fixed length sliding window of 2.56 sec with 50% overlap. The raw signal data per window were filtered using a median filter and a 3rd-order low-pass butterworth filter of $20Hz$ corner frequency. The jerk of the angular velocity, body and gravity acceleration were derived before computing standard statistical measures described in Table I and demonstrated in [23]. Overall, 561 features were extracted per window.

TABLE I
TABLE SHOWING THE LIST OF MEASUREMENTS FOR THE COMPUTATION OF FEATURE VECTORS ADAPTED FROM [23].

Function	Description
mean	Arithmetic mean
std	Standard deviation
mad	Median absolute deviation
max	Largest value in array
min	Smallest value in array
skewness	Frequency signal skewness
kurtosis	Frequency signal kurtosis
maxFreqInd	Largest frequency component
energy	Average sum of the squares
sma	Signal magnitude area
entropy	Signal entropy
iqr	Interquartile range
autoregression	4th order Burg autoregression coefficients
correlation	Pearson correlation coefficient
meanFreq	Frequency signal weighted average
energyBand	Spectral energy of a frequency band
angle	Angle between signal mean and vector

B. SVM and HMM

The original SVM developed in 1990s is a binary classification method [44]. Later, two strategies were developed to extend SVM in multi-class problems: 1) *one-against-all* strategy, which uses one SVM for each class and 2) *one-against-one* strategy, which uses a SVM for each pair of classes. Here the *one-against-all* strategy is used, which has shown superiority for multi-class classification problems [16], [45]. As described in [10], [46] the one-against-all approach consists of constructing k SVM models where k is the number of models. The ith model is then trained with all data samples belonging to class i as positive points and all other samples as negative points. Consequently, the classification of new instances are formulated using a winner-takes-all scheme given by equation 1, where f_i represents the ith classifier.

$$f(x) = arg\,max_i\, f_i(x) \tag{1}$$

The multi-class SVM is trained with a 561 (dimension) feature vector extracted from the measurements of the accelerometer and gyroscope sensors.

Hybrid SVM-HMM models have been shown to significantly improve classification accuracies over the standard SVM models [39]. The standard SVM is a discriminative classifier that does not provide class probabilities used by the HMM. However, simple post processing is proposed in [47] that can map the output of SVM to posterior class probabilities. The proposed method in [47] uses a sigmoid function to estimate these probabilities:

$$\hat{p}(x = m|f(\mathbf{y})) = (1 + exp(A_m f(\mathbf{y}) + B_m))^{-1} \tag{2}$$

where $f(\mathbf{y})$ is the output decision value of the SVM trained to separate class m from all other classes.

HMM is a basic approach for modeling correlated time series. The first order HMM is graphically shown in Fig. 1. It consists of a hidden state sequence $\{x_0, \cdots, x_k, \cdots\}$ and an observation sequence $\{\mathbf{y}_1, \cdots, \mathbf{y}_k, \cdots\}$. The observation vector $\mathbf{y}_k$ at time k corresponds to quantities that can be directly measured by sensors or computed from sensor output deterministically. The state variable x_k where $k \in \{1, \cdots, K\}$, represents the class label at time k that should be inferred. Note that this formulation is only valid for our problem as it does not depict the general formulation of the HMM. Temporal dependencies between class labels can be effectively modeled using HMM by assuming that the label

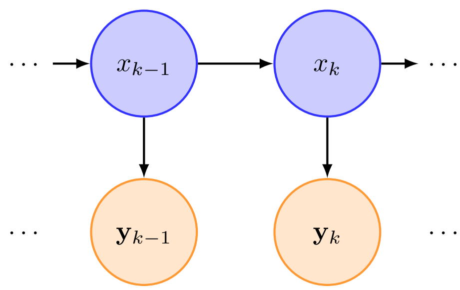

Fig. 1. Hidden Markov Model.

at time k is independent of the whole history of the process given the immediate previous label at time $k-1$.

HMM is characterized with two conditional probability density functions: 1) $p(x_k|x_{k-1})$ depicted as horizontal arrows in Fig. 1, which is referred to as the state transition model and 2) $p(\mathbf{y}_k|x_k)$ represented with vertical arrows in Fig. 1, which is called the emission model. The transition model is a categorical distribution

$$p(x_k|x_{k-1}=n) = Cat(\boldsymbol{\pi}_n), \tag{3}$$

where $\boldsymbol{\pi}_n = [\pi_{1,n}, \cdots, \pi_{K,n}]$ is the parameter vector of length K whose mth element $\pi_{m,n}$ equals the transition probability from state n to state m in subsequent time instances, i.e. $p(x_k = m|x_{k-1} = n)$. In total, there are K parameter vectors $\{\boldsymbol{\pi}_1, \cdots, \boldsymbol{\pi}_K\}$, each of which corresponds to one state label. If sequences of the class labels (hidden states) are available as training data, the transition parameters can be estimated using Maximum A-Posteriori (MAP) estimation by assuming a Dirichlet distribution prior for parameter vector

$$\boldsymbol{\pi}_n \sim Dir(\alpha), \tag{4}$$

where α, the concentration parameter, is set to 0.05 in this experiment. With this setup, the transition model MAP parameters can be calculated as

$$\hat{\pi}_{n,m} = \frac{\alpha + N_{n,m}}{K \times \alpha + \sum_{i=1}^{K} N_{n,i}}, \tag{5}$$

where $N_{n,m}$ is the number of times a transition from state n to state m occurs in the training sequences.

On the other hand, the emission model $p(\mathbf{y}_k|x_k)$ can be any kind of density function depending on the problem. For the HMM model, the observation vector would be the 561 feature vector but, defining a density function for such a high dimensional vector is unfeasible. However, using a well trained SVM classifier, it is still possible to calculate the posterior probabilities. The filtering task in the HMM is defined as the calculation of the posterior $p(x_k|\mathbf{y}_1, \cdots, \mathbf{y}_k, \cdots)$, which is done either by calculating forward filtering $\rho_(x_k) = p(x_k|\mathbf{y}_1, \cdots, \mathbf{y}_k)$ or T-lag forward-backward filtering $\rho_T(x_k) = p(x_k|\mathbf{y}_1, \cdots, \mathbf{y}_k, \cdots, \mathbf{y}_{k+T})$. These quantities can be calculated using forward $\alpha_k(m)$ and backward $\beta_k(m)$ values:

$$\rho(m) \propto \alpha_k(m), \tag{6}$$

$$\rho_T(m) \propto \alpha_k(m)\beta_k(m). \tag{7}$$

The forward value is calculated recursively as

$$\begin{aligned} \alpha_1(m) &= \gamma_1(m)/K, \\ \alpha_{k+1}(m) &= \gamma_{k+1}(m) \sum_{n=1}^{K} \alpha_k(n)\pi_{n,m}, \end{aligned} \tag{8}$$

and the backward value is calculated as

$$\begin{aligned} \beta_{k+T}(m) &= 1, \\ \beta_k(m) &= \sum_{n=1}^{K} \beta_{k+1}(n)\pi_{m,n}\gamma_{k+1}(n). \end{aligned} \tag{9}$$

where $\gamma_k(m)$ is the conditional probability of the label at time k given the feature vector calculated by (2) and m, n are hidden states x_k and x_{k+1} respectively. Having calculated the posterior probabilities, the activity class is found as the Maximum A Posteriori (MAP) state.

C. Artificial Neural Networks

Artificial Neural Networks (ANNs) is a machine learning paradigm, inspired by the way in which biological neural structures in the human brain, process information. Figure 2 shows the simplest model of an artificial neuron.

A single output (y) of the neuron is given by

$$y = f(\sum_i w_i x_i) = f(w^T x) \tag{10}$$

where x represents the input vector, w, the weight vector denotes the efficiencies of the neurons' synapses and f is the activation function. An Artificial Neural Network (ANN) is a

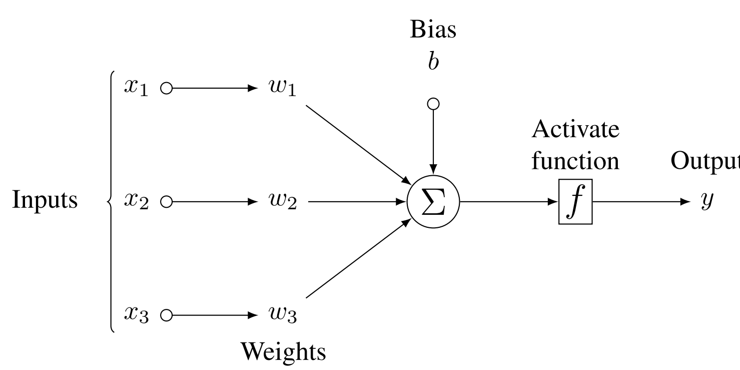

Fig. 2. Model of a basic artificial neuron.

network of neurons, which consists of an input vector, propagated via weights through the hidden layer until the activation reaches the output layer [48]. Figure 3 shows a generic ANN with 5 input units, 3 neurons in the hidden layer and 1 output. In this work, different ANN configurations including one and two hidden layers and varying number of neurons in the hidden

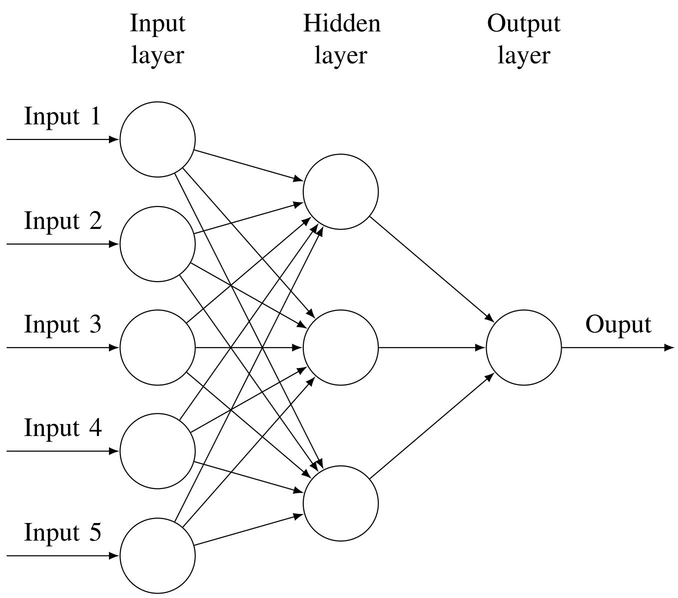

Fig. 3. Diagram showing the topology of the multi-layered ANN.

layers were evaluated. The configurations were trained using the scaled conjugate gradient algorithm [49] with the number of epochs tuned to 250. In the end, the one-hidden-layer ANN with 40 neurons gave the highest accuracy for our input vector.

IV. Results

As mentioned earlier, one of the primary goals of this work is to achieve the best classification performance for the development of a social connectedness application within the safety critical ambient assisted living domain. To determine the classification accuracy, the prominent K-fold cross-validation [50] (with $k = 10$) technique was applied to each classifier. This approach to validation was preferred to the traditional holdout method as it reduced the variance of the resulting estimates. The data was randomized and partitioned into ten equal parts, where 90% was used for training and 10% for testing in the cross-validation process. Consequently, every data point was used in the test set only once. The overall average performance was then calculated.

The confusion matrices of the classifiers are shown in Tables II, III and IV respectively. Rows of the confusion matrix represent the actual class while the columns represent the classifier output. The values in the confusion matrices are the number of instances in the test set. The last two columns of each table of the models show the classifiers' sensitivity and specificity scores for each class whilst the overall accuracies of the classifiers are provided at the bottom of the tables. Sensitivity, also known as true positive rate (TPR) or recall, defined by equation (11) estimates the probability of accurately identifying the class of a random data point.

$$TPR = \frac{True\,positives(T_p)}{T_p + False\,negatives(F_n)} \times 100 \qquad (11)$$

TABLE II
Confusion matrix using the SVM model.

	WALK	UP	DOWN	SIT	STAND	LAY	TPR(%)	TNR(%)
WALK	2703	11	20	0	0	0	98.9	99.7
UP	19	2353	18	3	3	6	98	99.7
DOWN	19	24	2289	0	2	2	98	99.7
SIT	1	0	0	2773	122	4	95.6	99
STAND	0	0	2	129	2804	0	95.5	99.0
LAY	0	5	0	0	0	2731	99.8	99.9

OVERALL ACCURACY 97.6%

TABLE III
Confusion matrix using the ANN model.

	WALK	UP	DOWN	SIT	STAND	LAY	TPR(%)	TNR(%)
WALK	2538	123	71	0	2	0	92.8	98.3
UP	111	2126	111	8	6	40	88.5	98.1
DOWN	108	104	2109	7	4	4	90.3	98.6
SIT	0	8	4	2540	304	44	87.6	97.7
STAND	3	5	7	281	2637	2	89.9	97.6
LAY	0	13	2	5	0	2716	99.3	99.3

OVERALL ACCURACY 91.4%

TABLE IV
Confusion matrix using the hybrid SVM-HMM model.

	WALK	UP	DOWN	SIT	STAND	LAY	TPR(%)	TNR(%)
WALK	2722	0	12	0	0	0	99.6	100
UP	0	2399	3	0	0	0	99.8	100
DOWN	0	0	2326	10	0	3	99.6	99.9
SIT	0	0	0	2892	8	15	99.7	99.9
STAND	0	0	0	10	2925	0	99.7	99.9
LAY	0	0	0	0	0	2736	100	100

OVERALL ACCURACY 99.7%

On the other hand, the specificity or the true negative rate (TNR), defined by equation (12), estimates the probability that a random data point not belonging to a class will be so rightfully identified by the classifier.

$$TNR = \frac{True\,negatives(T_n)}{T_n + False\,positives(F_p)} \times 100 \qquad (12)$$

Moreover, the overall accuracy (ACC) defined by equation (13) gives the fraction of data points correctly identified by the classifier.

$$ACC = \frac{T_p + T_n}{T_p F_p + T_n + F_n} \times 100 \quad (13)$$

All classes predicted by the stand-alone SVM achieved over 95% true positive and true negative rates while achieving an overall accuracy of 97.6%. From Table II, it was observed that dynamic activities i.e. walking, ascending and descending of stairs were a few times misclassified. Also, stationary activities i.e. standing, sitting and laying were occasionally indistinguishable by the SVM classifier .

The ANN classifier displayed the lowest detection accuracy among the classifiers with an overall accuracy of 91.4%. Like the SVM, walking, going up and downstairs, standing and sitting were occasionally indistinguishable. However, the ANN classifier displayed significantly more misclassifications when compared to the SVM. For example, the sensitivity (true positive rate) of sitting and standing were 87.6% and 89.9% respectively for ANN and, 95.6% and 95.5% respectively for SVM.

On the other hand, we noticed a very high overall accuracy of 99.7% for the SVM-HMM hybrid model outperforming the ANN and SVM models by 8.3% and 2.1% respectively. Notably, we observed improvements in predictions of all classes for the hybrid SVM-HMM classifier.

In sum, the hybrid SVM-HMM classification approach outperformed the other classifiers, showing an accuracy of 99.7%. Moreover, the performance of our SVM-HMM activity recognition model provides convincing evidence for its robustness and relevance in ambient assisted living environments.

V. Discussion and Conclusion

In this paper, we evaluated and compared three approaches to activity recognition (i.e. SVM, ANN and SVM-HMM) on real world data, *i. our own generated dataset* and *ii. a publicly available dataset*, using the smartphones' inertial sensors. This was done in an attempt to find the best classification accuracy for our bidirectional context awareness system within an AAL context. Experimental results reveal the superiority of the hybrid SVM-HMM classifier for human activity recognition against the ANN and SVM classifiers within an AAL context. Moreover, we have demonstrated the successful use of the smartphone for sensing in AAL, which made the data collection process inexpensive and easy to setup, and less obtrusive for our target users.

In total, we obtained the following recognition accuracies 91.4%, 97.6% and 99.7% for the ANN, SVM, SVM-HMM respectively. Albeit the SVM and ANN classifiers demonstrated good performance, the results show that SVM and ANN classifiers are less robust in dealing with the complexities and uncertainties of activity recognition data as stand alone classifiers. Through a combination of the time warping capabilities of the HMM and discriminative properties of the SVM, we obtained improved detection accuracies on all classes, demonstrating that the hybrid approach was better to overcome the HMM's weakness of discriminating between different classes.

Although smartphone based activity recognition offers great potential, it is not without limitations. For instance, within uncontrolled AAL environments, the smartphone's operational and functional challenges could impede large-scale adoption. Common challenges include limited battery life, memory capacity and processing power, privacy concerns and users deciding to turn off the device or possibly forgetting to charge or carry the device as discussed in [51], [52]. Moreover, location sensitivity could be an issue as the smartphone's sensors are heavily dependent on the sensor's positioning and orientation on the participant's body as posited in [53]. In the future, this could be addressed with the use of the magnetometer as proposed in [17].

On the other hand, the problem of participant sensitivity i.e., varying motion patterns among different people, makes recognition accuracy highly dependent on the participants used in the training and testing phases. By making our dataset publicly available, we contribute to the availability of diversified and reliable public mobile activity recognition datasets discussed in [51] through the inclusion of people from different countries with different age groups between 22 and 79 years.

While there is room for improvement in the area of smartphone based human activity recognition, we believe the application of machine learning to activity recognition problems using smartphone sensors in AAL holds promise for data scientists, designers and engineers to explore innovative solutions for the acquisition and presentation of activity data. As a continuation of this work, we will exploit our proposed hybrid approach in a social connectedness setting with an AAL context.

VI. Acknowledgments

This work was supported in part by the Erasmus Mundus Joint Doctorate (EMJD) in Interactive and Cognitive Environments (ICE), which is funded by Erasmus Mundus under the FPA no. 2010-2012. We would like to thank Mr. Sulman Baig for his assistance with this work and many thanks to our participants for volunteering in our activity recognition experiment.

References

[1] WHO, "World population ageing," *Technical Report, UN World Health Organization*, vol. 374, pp. 1–95, 2013.

[2] A. A. Helal, M. Mokhtari, and B. Abdulrazak, *The engineering handbook of smart technology for aging, disability, and independence*. Wiley Online Library, 2008.

[3] E. Y. Cornwell and L. J. Waite, "Social disconnectedness, perceived isolation, and health among older adults," *Journal of health and social behavior*, vol. 50, no. 1, pp. 31–48, 2009.

[4] A. Steptoe, A. Shankar, P. Demakakos, and J. Wardle, "Social isolation, loneliness, and all-cause mortality in older men and women," *Proceedings of the National Academy of Sciences*, vol. 110, no. 15, pp. 5797–5801, 2013.

[5] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hähnel, "Inferring activities from interactions with objects," *Pervasive Computing, IEEE*, vol. 3, no. 4, pp. 50–57, 2004.

[6] A. Bourke, J. O'brien, and G. Lyons, "Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm," *Gait & posture*, vol. 26, no. 2, pp. 194–199, 2007.

[7] B. Najafi, K. Aminian, A. Paraschiv-Ionescu, F. Loew, C. J. Büla, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly," *Biomedical Engineering, IEEE Transactions on*, vol. 50, no. 6, pp. 711–723, 2003.
[8] D. T. Van Bel, W. A. IJsselsteijn, and Y. A. de Kort, "Interpersonal connectedness: conceptualization and directions for a measurement instrument," in *CHI'08 extended abstracts on Human factors in computing systems*. ACM, 2008, pp. 3129–3134.
[9] S. Few, *Information dashboard design*. O'Reilly, 2006.
[10] J. L. R. Ortiz, *Smartphone-based human activity recognition*. Springer, 2015.
[11] N. A. Capela, E. D. Lemaire, and N. Baddour, "Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients," *PloS one*, vol. 10, no. 4, p. e0124414, 2015.
[12] L. Fiore, D. Fehr, R. Bodor, A. Drenner, G. Somasundaram, and N. Papanikolopoulos, "Multi-camera human activity monitoring," *Journal of Intelligent and Robotic Systems*, vol. 52, no. 1, pp. 5–43, 2008.
[13] V. Bastani, L. Marcenaro, and C. Regazzoni, "A particle filter based sequential trajectory classifier for behavior analysis in video surveillance," in *Image Processing (ICIP), 2015 IEEE International Conference on*, Sept 2015, pp. 3690–3694.
[14] S. Patel, H. Park, P. Bonato, L. Chan, M. Rodgers *et al.*, "A review of wearable sensors and systems with application in rehabilitation," *J Neuroeng Rehabil*, vol. 9, no. 12, pp. 1–17, 2012.
[15] J. M. Corchado, J. Bajo, D. Tapia, A. Abraham *et al.*, "Using heterogeneous wireless sensor networks in a telemonitoring system for healthcare," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 234–240, 2010.
[16] 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*. Springer, 2012, pp. 216–223.
[17] Y. E. Ustev, O. Durmaz Incel, and C. Ersoy, "User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal," in *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*. ACM, 2013, pp. 1427–1436.
[18] K. Davis, E. Owusu, C. Regazzoni, L. Marcenaro, L. Feijs, and J. Hu, "Perception of human activities a means to support connectedness between the elderly and their caregivers," in *Proceedings of the 1st International Conference on Information and Communication Technologies for Ageing Well and e-Health*. SCITEPRESS, 2015, pp. 194–199.
[19] K. Davis, J. Hu, L. Feijs, and E. Owusu, "Social hue: A subtle awareness system for connecting the elderly and their caregivers," in *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*. IEEE, 2015, pp. 178–183.
[20] E. D. Mynatt, J. Rowan, S. Craighill, and A. Jacobs, "Digital family portraits: supporting peace of mind for extended family members," in *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM, 2001, pp. 333–340.
[21] J. Rowan and E. D. Mynatt, "Digital family portrait field trial: Support for aging in place," pp. 521–530, 2005.
[22] M. Memon, S. R. Wagner, C. F. Pedersen, F. H. A. Beevi, and F. O. Hansen, "Ambient assisted living healthcare frameworks, platforms, standards, and quality attributes," *Sensors*, vol. 14, no. 3, pp. 4312–4341, 2014.
[23] J.-L. Reyes-Ortiz, L. Oneto, A. Samà, X. Parra, and D. Anguita, "Transition-aware human activity recognition using smartphones," *Neurocomputing*, vol. 171, pp. 754–767, 2016.
[24] 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.
[25] T. W. Liao, "Clustering of time series dataa survey," *Pattern recognition*, vol. 38, no. 11, pp. 1857–1874, 2005.
[26] D. Rodriguez-Martin, A. Sama, C. Perez-Lopez, A. Catala, J. Cabestany, and A. Rodriguez-Molinero, "Svm-based posture identification with a single waist-located triaxial accelerometer," *Expert Systems with Applications*, vol. 40, no. 18, pp. 7203–7211, 2013.
[27] L. Zhao, X. Wang, G. Sukthankar, and R. Sukthankar, "Motif discovery and feature selection for crf-based activity recognition," in *Pattern Recognition (ICPR), 2010 20th International Conference on*. IEEE, 2010, pp. 3826–3829.
[28] F. Paganelli and D. Giuli, "An ontology-based system for context-aware and configurable services to support home-based continuous care," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 2, pp. 324–333, 2011.
[29] P. Lubina and M. Rudzki, "Artificial neural networks in accelerometer-based human activity recognition," in *Mixed Design of Integrated Circuits & Systems (MIXDES), 2015 22nd International Conference*. IEEE, 2015, pp. 63–68.
[30] D. J. Patterson, D. Fox, H. Kautz, and M. Philipose, "Fine-grained activity recognition by aggregating abstract object usage," in *Wearable Computers, 2005. Proceedings. Ninth IEEE International Symposium on*. IEEE, 2005, pp. 44–51.
[31] T. V. Duong, H. H. Bui, D. Q. Phung, and S. Venkatesh, "Activity recognition and abnormality detection with the switching hidden semi-markov model," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 838–845.
[32] Y. Kwon, K. Kang, and C. Bae, "Unsupervised learning for human activity recognition using smartphone sensors," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6067–6074, 2014.
[33] A. Jurek, C. Nugent, Y. Bi, and S. Wu, "Clustering-based ensemble learning for activity recognition in smart homes," *Sensors*, vol. 14, no. 7, pp. 12 285–12 304, 2014.
[34] E. Kim, S. Helal, and D. Cook, "Human activity recognition and pattern discovery," *Pervasive Computing, IEEE*, vol. 9, no. 1, pp. 48–53, 2010.
[35] F. J. Ordónez, P. de Toledo, and A. Sanchis, "Activity recognition using hybrid generative/discriminative models on home environments using binary sensors," *Sensors*, vol. 13, no. 5, pp. 5460–5477, 2013.
[36] J. Suutala, S. Pirttikangas, and J. Röning, "Discriminative temporal smoothing for activity recognition from wearable sensors," in *Ubiquitous Computing Systems*. Springer, 2007, pp. 182–195.
[37] A. Ganapathiraju, J. E. Hamaker, and J. Picone, "Applications of support vector machines to speech recognition," *Signal Processing, IEEE Transactions on*, vol. 52, no. 8, pp. 2348–2355, 2004.
[38] Y.-L. Lin and G. Wei, "Speech emotion recognition based on hmm and svm," in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, vol. 8. IEEE, 2005, pp. 4898–4901.
[39] M. F. Valstar and M. Pantic, "Combined support vector machines and hidden markov models for modeling facial action temporal dynamics," in *Human–Computer Interaction*. Springer, 2007, pp. 118–127.
[40] J. Ye, S. Dobson, and S. McKeever, "Situation identification techniques in pervasive computing: A review," *Pervasive and mobile computing*, vol. 8, no. 1, pp. 36–66, 2012.
[41] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern recognition letters*, vol. 29, no. 16, pp. 2213–2220, 2008.
[42] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN*, 2013.
[43] (2016) Smartphone dataset for human activity recognition (har) in ambient assisted living (aal). [Online]. Available: https://github.com/kadiand/Smartphone-HAR-Data-For-AAL
[44] C. Cortes and V. Vapnik, "Support-vector networks," *Machine learning*, vol. 20, no. 3, pp. 273–297, 1995.
[45] R. Rifkin and A. Klautau, "In defense of one-vs-all classification," *The Journal of Machine Learning Research*, vol. 5, pp. 101–141, 2004.
[46] C.-W. Hsu and C.-J. Lin, "A comparison of methods for multiclass support vector machines," *Neural Networks, IEEE Transactions on*, vol. 13, no. 2, pp. 415–425, 2002.
[47] J. C. Platt, "Probabilities for sv machines," in *Advances in Large Margin Classifiers*. MIT Press, March 1999, pp. 61–74. [Online]. Available: http://research.microsoft.com/apps/pubs/default.aspx?id=69187
[48] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice Hall, 2007.
[49] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural networks*, vol. 6, no. 4, pp. 525–533, 1993.
[50] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *AAAI*, vol. 5, 2005, pp. 1541–1546.
[51] D. Choujaa and N. Dulay, "Activity recognition from mobile phone data: State of the art, prospects and open problems," *Imperial College London*, 2009.

[52] N. Eagle and A. Pentland, "Reality mining: sensing complex social systems," *Personal and ubiquitous computing*, vol. 10, no. 4, pp. 255–268, 2006.
[53] X. Su, H. Tong, and P. Ji, "Activity recognition with smartphone sensors," *Tsinghua Science and Technology*, vol. 19, no. 3, pp. 235–249, 2014.

View publication stats