

# Robust Human Intensity-Varying Activity Recognition using Stochastic Approximation in Wearable Sensors

Nabil Alshurafa, Wenyao Xu, Jason J. Liu,  
Ming-Chun Huang, Bobak Mortazavi, Majid Sarrafzadeh  
Wireless Health Institute, Department of Computer Science  
University of California, Los Angeles  
Email: {nabil, wenyao, jasonliu, mingchuh, bobakm, majid}@cs.ucla.edu

Christian Roberts  
School of Nursing  
Exercise and Metabolic Disease Research Laboratory  
University of California, Los Angeles  
Email: {croberts}@sonnet.ucla.edu

**Abstract**—Detecting human activity independent of intensity is essential in many applications, primarily in calculating metabolic equivalent rates (MET) and extracting human context awareness from on-body inertial sensors. Many classifiers that train on an activity at a subset of intensity levels fail to classify the same activity at other intensity levels. This demonstrates weakness in the underlying activity model. Training a classifier for an activity at every intensity level is also not practical. In this paper we tackle a novel intensity-independent activity recognition application where the class labels exhibit large variability, the data is of high dimensionality, and clustering algorithms are necessary. We propose a new robust Stochastic Approximation framework for enhanced classification of such data. Experiments are reported for each dataset using two clustering techniques, K-Means and Gaussian Mixture Models. The Stochastic Approximation algorithm consistently outperforms other well-known classification schemes which validates the use of our proposed clustered data representation.

**Keywords**—Stochastic Approximation Model; Classification; Mixture Models; Clustering; Intensity-Varying Activity;

## I. INTRODUCTION

Wearable sensors, specifically inertial sensors, continue to be used in activity recognition systems and devices [1]. For example, volumes of literature have shown the value of using accelerometer-based activity monitors [2], [3]. However, most research either trains and tests activity at a single intensity level [4], or omits the intensity of the activity all together [5], [6]. Features collected on an activity categorized as running or walking will not only vary from one subject to the next, but will also vary with intensity. Yet when we physically see someone running or walking at any intensity level we can readily distinguish between the two, even though our eyes haven't seen running and walking at every possible intensity level. The challenge remains to train a classifier on a subset of activity intensity levels, and test on different intensity levels. Only when a classifier performs well in classifying an activity at new intensity levels can it be claimed to be robust, providing intensity-independent activity classification.

Due to the large variance and overlap in the feature sets, it is quite challenging for a classification algorithm trained on one intensity level to distinguish walking and running at another intensity level [7]. Due to large inter-subject variability, the

nature of the data collected from an accelerometer at multiple intensities will be scattered with high variance, exhibiting overlapping categories across intensity levels. Figure 1 provides an illustration of the large variance in the features collected across subjects and intensity levels, where someone walking at 2mph may present similar characteristics to someone walking at 4mph.

Class models are typically generated using raw sample points within each class label (activity category), resulting in complex and computationally inefficient systems. Some use clustering to enhance computational efficiency. We attempt to generate an activity model that uses clustering combined with uncertainty within the clusters to learn a robust activity model that can classify the same activity at a range of intensity levels, resulting in a more efficient formulation. Our framework is compared against other well-known classification schemes such as SVM, kNN, and C4.5 Decision Trees.

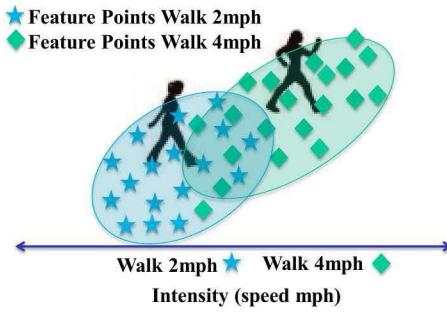


Fig. 1. Overlap among features when walking at varying intensities

In many health-related fields, interest in energy expenditure has sparked the need to analyze the intensity of human activity. Ainsworth et al. attempted to compile a compendium using regression models to map common physical activities to Metabolic Equivalent of Tasks (METs) [8]. However, Kozey et al. [9], show that linear regression models are inappropriate for accurately predicting METs from accelerometer data. Albinali et al. [4] argue that automatic detection of physical activity type prior to using activity regressions enhances the estimate of energy expenditure. Moreover, activity monitors that rely on such equations have no physically interpretable meaning of the activity being performed. In this paper we focus on identifying

a model for a physical activity type that is independent of intensity, which we believe will aid in understanding the underlying nature of a physical activity type.

Vathsangam et al. [10] analyze walking at only two intensity levels, predicting energy expenditure from treadmill walking using hip-worn inertial sensors; applying regression techniques such as Least-Squares Regression (LSR), Bayesian Linear Regression (BLR) and Gaussian Process Regression (GPR), without classifying the activities. Ravi et al. [5] and Kwapisz et al. [11] reach subject-dependent classification accuracies above 90%, however they did not focus on distinguishing the intensity of the activity. Tapia et al. [7] attempted to recognize physical activities and their intensity levels using a C4.5 Decision Tree classifier, yet achieved poor subject-independent accuracies of 58.2%, and also did not train on a subset of activity intensity levels and test on another intensity.

Miller et al. [12] has shown success using mixture models and EM-based algorithms for new class discovery with mixed labeled and unlabeled data sets. Sun et al. [13] attempt to classify EEG signals using Stochastic Approximation to update parameters of Bayesian classifiers on EEG signals. Compared to existing studies, our work differs accordingly:

- 1) Intensity Independent classification: Our model-based approach attempts to represent activities holistically, and is capable of inferring other intensity levels from a subset of intensity levels.
- 2) Stochastic Approximation framework: We utilize clustering algorithms and Stochastic Approximation to create a model for each activity, and then use a Stochastic Approximation nearest-neighbor heuristic for classification.

This paper is organized as follows. In Section II we briefly review background on clustering algorithms and Stochastic Approximation. Then we describe our intensity-independent activity recognition framework in Section III. We present our experimental results and analysis in Section IV. Finally, we conclude and discuss future work in Section VI.

## II. PRELIMINARY

### A. Clustering and data representation

Clustering has been widely used in pattern recognition applications [14] as a means of organizing data into groups or clusters based on a similarity metric. It is mainly applied to discover classes and structures of data in an unsupervised-based learning fashion. It is also an efficient method for representing data, allowing for fast retrieval, and dimensionality reduction. In this section we mainly cover two types of clustering algorithms that are further examined in the remainder of the paper: K-Means exclusive clustering, and a Gaussian Mixture Model (GMM) based probabilistic clustering algorithm.

Let us assume we have  $n$  sample points labeled  $x_i \in R^m$ . Each  $x_i$  represents a vector of features collected on a sample. The K-Means clustering technique is one of the simplest and most popular unsupervised learning algorithms [15]. It attempts to find  $k$  clusters that partition the sample points. With initial random placement of  $k$  centroids  $\mu_j \in R^m$ , each point  $x_i$  is assigned to the closest centroid. Where  $x_i^j$  is now the data

point that is assigned to cluster  $j$ .  $k$  centroids are iteratively recalculated and points reassigned such that the following sum of squared distances to the cluster centers is minimized:

$$J(\mu_1, \mu_2, \dots, \mu_k) = \sum_{j=1}^k \sum_{i=1}^n \|x_i^j - \mu_j\|^2 \quad (1)$$

Training samples are tightly clustered around the centroids to serve as a compact representation for the training data. K-Means is simple to implement, but does not necessarily find a global optimal solution based on the objective function. K-Means is also known to be sensitive to the initial randomly selected cluster centers.

The Gaussian Mixture Model (GMM) [16] is another widely used, model-based clustering technique. GMM's optimize the fit between data and a parametric distribution (like a Gaussian or Poisson distribution), where the entire data is modeled by a mixture of such distributions. Each category of activity  $A_i$  can be clustered into a set of learned Gaussian distributions where  $\omega_j$  is the prior probability of the  $j^{th}$  distribution, with mean  $\mu_j$ , variance  $\sigma_j^2$ , and is represented by  $N(\mu_j, \sigma_j^2)$ . The probability  $p(x_i)$  that a data point  $x_i$  belongs to a category  $A_i$  can be represented by a set of mixture models with varying distributions, and is defined by:

$$p(x_i) = \sum_{j=1}^k \omega_j N(x_i | \mu_j, \sigma_j^2) \quad (2)$$

Given a set of points  $X = < x_1, x_2, \dots, x_n >$ , drawn from an unknown distribution, we estimate the parameters of the GMM model  $\theta$  that fits the data (i.e.  $\omega_j$ ,  $N(\mu_j, \sigma_j^2)$ ). The solution is to maximize the likelihood  $p(X|\theta)$  of the data with regards to the model parameters:

$$\arg \max_{\theta} p(X|\theta) = \arg \max_{\theta} \prod_{i=1}^N p(x_i|\theta) \quad (3)$$

The Expectation Maximization algorithm is used in practice to find the mixture of Gaussians that can model the data [16].

### B. Stochastic Approximation

Stochastic Approximation is a robust method used to solve objective functions while taking into account uncertainty or possible variation in the data [17]. It typically involves a non-tractable convex optimization problem, where the objective or its derivatives are difficult to evaluate.

We consider a matrix  $A \in R^{m \times n}$  with the basic objective  $\|Ax - b\|$ , but also wish to take into account uncertainty. In many cases the number of unknowns is larger than the feature space,  $n > m$ , making it an underdetermined system. In our application, we use clustering algorithms to ensure that  $n < m$ , resulting in an overdetermined system. The mean of  $A$  is  $\bar{A}$ , so it can be described as:

$$A = \bar{A} + U \quad (4)$$

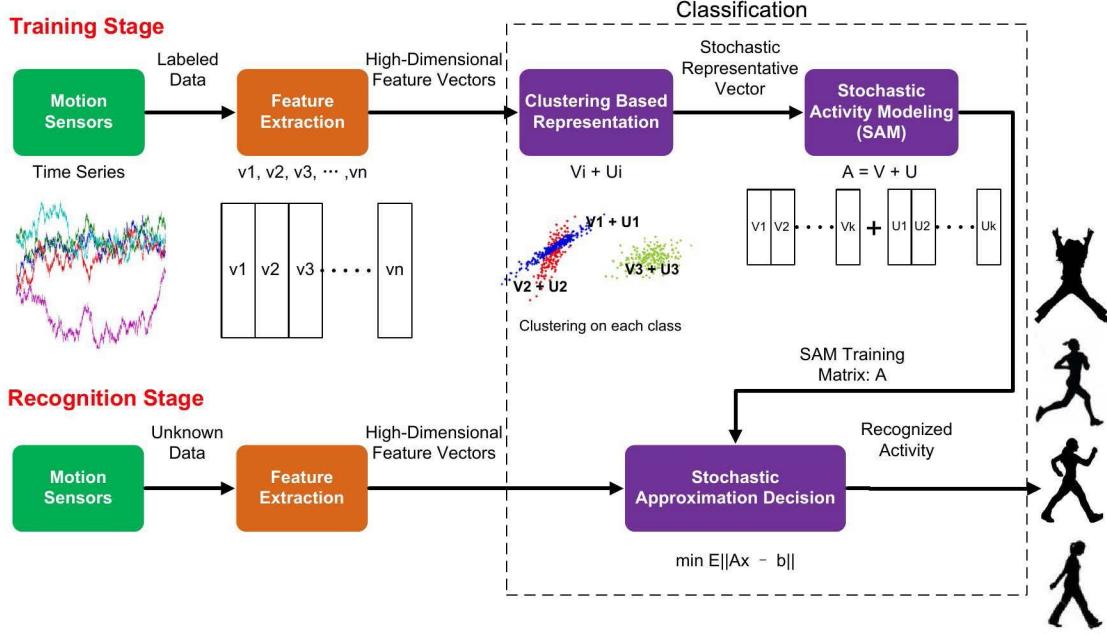


Fig. 2. Stochastic Approximation framework

where  $U$  is a random matrix with zero mean.  $U$  describes the statistical variation of  $A$  and can be used to model uncertainty or large variation in the data matrix  $A$ . A natural objective is to minimize the expected value of  $Ax - b$ :

$$\min_{x \in R^n} E\|Ax - b\| \quad (5)$$

This is a general form of the stochastic robust approximation problem. Some variations of the statistical robust approximation problem are tractable, for example the statistical robust least squares problem, where we minimize the Euclidean norm of  $Ax - b$ :

$$\min_{x \in R^n} E\|Ax - b\|_2^2 \quad (6)$$

This objective function has a closed form solution that can be expressed as follows:

$$\begin{aligned} E\|Ax - b\|_2^2 &= E(\bar{A} - b + Ux)^T(\bar{A}x - b + Ux) \\ &= (\bar{A}x - b)^T(\bar{A}x - b) + Ex^T U^T Ux \\ &= \|\bar{A}x - b\|_2^2 + x^T Px \end{aligned} \quad (7)$$

where  $P = E(U^T U)$ . This is exactly the form of a regularized least-squares problem [17]:

$$\min_{x \in R^n} \|\bar{A}x - b\|_2^2 + \|P^T x\|_2^2 \quad (8)$$

And the solution is:

$$x = (\bar{A}^T \bar{A} + P^{-1})^{-1} \bar{A}^T b \quad (9)$$

### III. OUR FRAMEWORK

In this section we present our framework for intensity independent classification of activity. Figure 2 illustrates our framework comprising four important components: feature extraction, clustering based representation, Stochastic Activity Modeling (SAM), and Stochastic Approximation decision classification. The training phase begins by extracting features from the accelerometer time series data, which is then fed into a clustering based algorithm. The data is then grouped by activity into clusters, and the clusters' mean and variance are then combined to form a corresponding SAM. In the recognition phase, the sample test points are then compared against each activity category's SAM in a stochastic decision making component. We describe the details of these components in this section.

#### A. Feature Extraction

There are several studies that analyze varying features that are best utilized for human activity based on accelerometer data. Table I lists the main features that have been shown to be useful in classifying activity [5], [18], [19], [20]. Using each  $X, Y, Z$  acceleration axis generates a total of 45 features per segment, where a segment is a fixed time subdivision of the accelerometer data. Each segment results in a feature vector that belongs to one of  $N$  categories:  $A_1, A_2, \dots, A_N$ , where each activity category represents a specific activity type.

TABLE I. FEATURE TABLE

Mean	Standard Deviation	Mean Derivatives
Median	Pairwise Correlation	Interquartile Range
Skewness	Root Mean Square	Zero Crossing Rate
Variance	Mean Crossing Rate	Kurtosis

## B. Applying Clustering Algorithms

In order to solve the least squares problem defined in Equation 8, we need to obtain the matrix  $\bar{A}_i \in R^{m \times k}$  for each activity. We collect all the samples from class  $A_i$  into  $k$  clusters  $X_1, X_2, \dots, X_k$  using K-Means or the GMM clustering algorithm. We ensure that the generated number of clusters  $k$  is less than the number of features  $m$  in order to make sure we have an overdetermined system. Each cluster  $X_i$  can be represented by  $V_i + U_i$ , where  $V_i$  is the cluster center, and  $U_i$  is the random noise based on the distribution of cluster  $X_i$ .

We tested both the K-Means and GMM clustering methods. From each generated cluster we extract the mean and variance, and combine them for each activity category  $A_i$  to form the  $V_i$  and  $U_i$  matrix, respectively.

## C. Stochastic Activity Modeling and Classification

Once the data is clustered, we then combine all the clusters of each activity to form a SAM matrix  $B_i$  for each activity  $A_i$ . Each category contains  $k$  clusters which are generated from  $n_i$  training samples, each having  $m$  features.

$$\begin{aligned} B_i &= [V_1, V_2, \dots, V_i, \dots, V_k] + [U_1, U_2, \dots, U_i, \dots, U_k] \\ &= \begin{bmatrix} v_{11} & \dots & v_{1k} \\ v_{21} & \dots & v_{2k} \\ \vdots & \ddots & \vdots \\ v_{m1} & \dots & v_{mk} \end{bmatrix} + \begin{bmatrix} u_{11} & \dots & u_{1k} \\ u_{21} & \dots & u_{2k} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mk} \end{bmatrix} \quad (10) \end{aligned}$$

where  $v_{i,j}$  represents the center of the  $i^{th}$  feature of the  $j^{th}$  cluster, and  $u_{i,j}$  represents the variance of the  $i^{th}$  feature of the  $j^{th}$  cluster. Given segment  $b$ , we want to make a decision as to which SAM  $B_i$  belongs. To do this we solve the following using Equation 8:

$$\min_{i=1 \dots N} E \|B_i x - b\|_2^2 \quad (11)$$

In other words we iterate through each of the  $N$  SAMs and determine to which activity category  $b$  belongs by choosing the one with the minimum reconstruction error, which is similar to searching for the nearest-neighbor SAM. We validate the classification performance on collected data in an in-lab setting.

## IV. EXPERIMENTS AND EVALUATION

Our experiment involved participants performing a predefined set of activities. Prior to performing any activity, the subject would wear a belt-like strap around the waist. A Gulf Coast Data Concepts X6-2 mini tri-axis accelerometer is embedded inside the belt, positioned on the right-side of the participants waist. It is capable of measuring acceleration with 12-bit resolution, setting the detection range at  $\pm 2g$ . The accelerometer was set to record readings at a frequency of 80Hz. Features in Table I were extracted from the accelerometer. In order to clearly identify the intensity level of walking and running, the experiments involved performing activity on a treadmill.

## A. Data Collection

We collected data from twelve subjects, ranging in age from 20 to 28 years. Each subject wore the accelerometer embedded strap around their waist and performed several exercises on a Merit Fitness 715T Plus treadmill: walking at three intensity levels, running at three intensity levels, and jumping jacks were performed to the side of the treadmill. Each activity was performed for five minutes, with five minutes of rest between each activity. When we were analyzing results from generic classifiers, we found the majority of misclassifications occurred between the running and walking activity types. For this reason our intensity independent model targets correct classification of walking and running. Table II provides the labels for each category.

TABLE II. ACTIVITY TYPES

W1	Walking 2.5mph	R1	Running 4mph
W2	Walking 3.5mph	R2	Running 5mph
W3	Walking 5mph	R3	Running 6mph

## B. Classification Performance and Evaluation

In order to discover the potential of our Stochastic Approximation framework, we compared our classifier against other well-known classifiers. Each classifier was trained with eleven subjects and tested on one, using six class labels (or activity categories), performing Leave One Out Cross Validation by subject. Figure 3 shows the results for each classifier. Using the SVM, kNN and C4.5 Decision Tree (DT) classification algorithms we achieved F-measures (the harmonic mean of precision and recall) of 81.5%, 82.4% and 75.0% respectively. Our Stochastic Approximation framework with the GMM clustering algorithm yielded a 94% precision, 90% recall, and a 92% F-measure. Under the K-Means clustering algorithm, the Stochastic Approximation classifier resulted in an F-measure of 85%, which remains superior to the other classification techniques. Based on these results we conclude that the Stochastic Approximation classification framework is better able to handle the large variation across activity types.

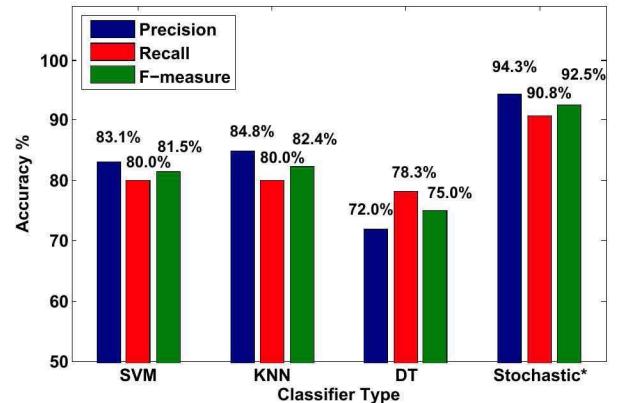


Fig. 3. The precision, recall and F-measure across multiple classifiers applied to walking and running at multiple intensity levels. The Stochastic Approximation Classifier with GMM clustering outperforms SVM, kNN, and C4.5 Decision Tree (DT).

We also evaluate the stability of our classification algorithm by analyzing the variation during cross validation. Stability

is another important measure that describes how closely a classifier evaluates results if given different data. Figure 4 shows the classification variation for each of the classifiers. We can conclude that the Stochastic Approximation framework not only outperforms other classifiers in precision, recall and F-measure results, but also produces stable results across multiple runs, resulting in low variance compared to other classifiers.

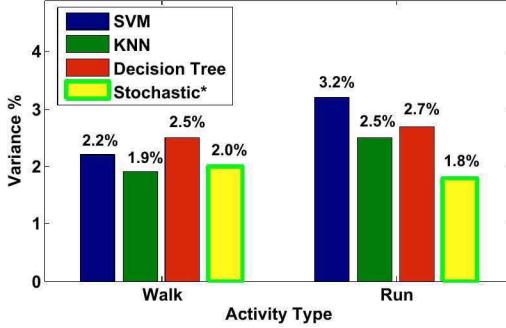


Fig. 4. The variance across classifiers during cross validation shows that the Stochastic Approximation Classifier is a robust classifier with low variance.

### C. Stochastic nature of human activity

When we separated each intensity level and activity type into its own category, training the classifier on six categories, the classifiers were trained on each intensity level, performing well on the test data. Table III shows high average precision and recall. While the classifier performs sufficiently well when trained on six categories, our primary goal has been to demonstrate the classifiers generalizability when trained on a subset of intensity levels. Table IV shows the results of nine classifiers, each trained on a different subset of intensity levels. We show the average category F-measure, which is the average F-measure across intensity levels ( $Walk_A$  column is an average of W1, W2, and W3, and the  $Run_A$  column is an average of R1, R2, and R3). Table IV also shows the F-measure for the specific intensity independent activity excluded from the training set, e.g. the first row trained on W1, W2 and R1, R2 and tested on W3 and R3, as shown in the  $Walk_T$  and  $Run_T$  columns in Table IV.

It is interesting to note that of the nine classifiers the one that performs the best is the one that trains on the extremes of each category, W1, W3 and R1, R3. This result is illustrated in the fifth row of Table IV, where the Stochastic Approximation Classifier is capable of extrapolating information of W2 from W1 and W3, and R2 from R1 and R3. Typically when a subject runs at 2mph, they are not really running at 2mph, but fluctuate in speeds slightly above and below 2mph, and this noise is represented in the SAM of the activity, since it takes into account the variance in the clusters. The SAM is capable of extrapolating information about adjacent intensity levels from the variance of the clustered data. This supports the belief that in order to train a classifier for a given physical activity, we may only need to train the Stochastic Approximation Classifier on a subset of intensity levels.

The classifier with the worst performance was the one trained on W1, W2 and R1, R3, shown in row 2 of Table IV. This is due to the fact that there is much overlap between W3 and R2, since they represent walking and running at 5mph,

respectively. Testing on the run category R2 yields a slightly higher F-measure, 80.5% compared to 79.4% for W3. This shows the classifier is able to perform a little better on R2 given it was trained on R1 and R3, however the higher intensity activity, W3, is more challenging to classify given the training of lesser intensity activities W1 and W2.

We wanted to investigate the choice of clustering algorithm on the performance of our Stochastic Approximation framework. We compared the outcomes of two clustering techniques, K-Means and GMM. It is interesting to note that in our application using GMM outperforms K-Means on average. This happens because K-Means essentially represents each category by a cluster of centroids, while GMM represents each category by a cluster of distributions, each with a mean, variance and probability. Therefore, GMMs can better represent the model when the data is sparse with high variability. Table V provides the details of this comparison.

TABLE V. COMPARING K-MEANS AND GMM CLUSTERING ALGORITHMS

K-Means		GMM	
Precision	Recall	Precision	Recall
89.2%	86.3%	94.4%	93.7%

## V. CONCLUSION

Tri-axis accelerometers are increasingly prevalent in wearable sensors to derive behavioral patterns and context from human motion. Subject independent activity classification using such accelerometers is critical in developing systems that are robust and scalable, and perform well in a real-world setting. However, for activities with varying intensity levels, one needs to also analyze the ability of a system to detect activity independent of intensity. We show promise in our Stochastic Approximation framework in its ability to extrapolate unknown intensity levels from a few known intensity levels. Our framework is capable of outperforming other well-known algorithms, and also performs better using a GMM clustering algorithm as opposed to a K-Means clustering algorithm. Our system will be useful in many healthcare applications including calculating energy expenditure and context awareness, as well as accurate classification in an activity-based environment.

In the future, we can attempt to test our system on other intensity varying activities, such as bicycling and across more subjects. We would also like to test our classifier with other clustering algorithms: HMM and Fuzzy C-means along with hierarchical clustering. We also hope to test our framework on individuals running and walking at a greater range of intensity levels.

## REFERENCES

- [1] S. Patel, H. Park, P. Bonato, L. Chan, and M. Rodgers, "A review of wearable sensors and systems with application in rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 9, no. 1, p. 21, 2012.
- [2] P. S. Freedson, E. Melanson, and J. Sirard, "Calibration of the Computer Science and Applications, Inc. accelerometer," *Med Sci Sports Exerc*, vol. 30, no. 5, pp. 777–781, May 1998.
- [3] P. S. Freedson, K. Lyden, S. Kozey-Keadle, and J. Staudenmayer, "Evaluation of artificial neural network algorithms for predicting METs and activity type from accelerometer data: validation on an independent sample," *J. Appl. Physiol.*, vol. 111, no. 6, pp. 1804–1812, Dec 2011.

TABLE III. CONFUSION MATRIX FOR THE STOCHASTIC APPROXIMATION CLASSIFIER TRAINING ON SIX CATEGORIES

		Walk			Run					Category Recall
Category	Activity	W1	W2	W3	R1	R2	R3	Total	Recall	Category Recall
Walk	W1	75	5	1	0	0	0	81	92.6%	93.9%
	W2	7	71	1	2	2	0	83	85.5%	
	W3	3	6	76	5	5	2	97	78.4%	
Run	R1	1	2	2	73	2	2	82	89.0%	95.6%
	R2	0	1	4	3	74	6	88	84.1%	
	R3	0	0	1	1	2	76	81	93.8%	
	Total	86	85	85	85	85	85			
Precision	87.2%	83.5%	89.4%	85.9%	87.1%	89.4%				
Category Precision	95.7%			93.7%						

TABLE IV. CLASSIFICATION RESULTS FOR THE STOCHASTIC APPROXIMATION CLASSIFIER TRAINING ON A SUBSET OF CATEGORIES

Classifier Number	Training (X means selected in training)						Testing			
	W1	W2	W3	R1	R2	R3	Average Category F-measure	Intensity Independent F-measure		
							$Walk_A^*$	$Run_A^*$	$Walk_T^\dagger$	$Run_T^\dagger$
1	X	X		X	X		84.5%	85.5%	82.1% (W3)	83.2% (R3)
2	X	X		X		X	81.5%	83.5%	79.4% (W3)	80.5% (R2)
3	X	X			X	X	86.5%	84.5%	82.8% (W3)	80.9% (R1)
4	X		X	X	X		89.0%	88.5%	83.6% (W2)	83.2% (R3)
5	X		X	X		X	94.0%	92.5%	84.1% (W2)	85.6% (R2)
6	X		X		X	X	86.5%	85.5%	83.4% (W2)	82.1% (R1)
7		X	X	X	X		89.5%	89.0%	85.1% (W1)	84.3% (R3)
8		X	X	X		X	90.5%	91.5%	84.2% (W1)	83.2% (R2)
9		X	X		X	X	86.5%	85.5%	79.7% (W1)	80.1% (R1)

\* $Walk_A$  and  $Run_A$  represent the average F-measure across all intensity values.

† $Walk_T$  and  $Run_T$  represent the F-measure of the tested intensity independent activity not used in the training set.

- [4] F. Albinali, S. Intille, W. Haskell, and M. Rosenberger, "Using wearable activity type detection to improve physical activity energy expenditure estimation," in *Proceedings of the 12th ACM international conference on Ubiquitous computing*, ser. Ubicomp '10. ACM, 2010, pp. 311–320.
- [5] N. Ravi, N. D. P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence(IAAI)*. AAAI Press, 2005, pp. 1541–1546.
- [6] J. Yang, "Toward physical activity diary: motion recognition using simple acceleration features with mobile phones," in *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics*, ser. IMCE '09, 2009, pp. 1–10.
- [7] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman, "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart monitor," in *IN: PROC. INT. SYMP. ON WEARABLE COMP*, 2007.
- [8] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett, C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, and A. S. Leon, "Compendium of Physical Activities: a second update of codes and MET values," *Med Sci Sports Exerc*, vol. 43, no. 8, pp. 1575–1581, Aug 2011.
- [9] S. L. Kozey, K. Lyden, C. A. Howe, J. W. Staudenmayer, and P. S. Freedson, "Accelerometer output and MET values of common physical activities," *Med Sci Sports Exerc*, vol. 42, no. 9, pp. 1776–1784, Sep 2010.
- [10] H. Vathsangam, A. Emken, E. Schroeder, D. Spruijt-Metz, and G. Sukhatme, "Determining energy expenditure from treadmill walking using hip-worn inertial sensors: An experimental study," *Biomedical Engineering, IEEE Transactions on*, vol. 58, no. 10, pp. 2804 –2815, Oct 2011.
- [11] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, Mar. 2011.
- [12] D. J. Miller and J. Browning, "A mixture model and em-based algorithm for class discovery, robust classification, and outlier rejection in mixed labeled/unlabeled data sets," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 11, pp. 1468–1483, 2003.
- [13] S. Sun, M. Lan, and Y. Lu, "Adaptive eeg signal classification using stochastic approximation methods," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, 31 2008–april 4 2008, pp. 413–416.
- [14] V. Estivill-Castro, "Why so many clustering algorithms: a position paper," *SIGKDD Explor. Newsl.*, vol. 4, no. 1, pp. 65–75, Jun. 2002.
- [15] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1967, pp. 281–297.
- [16] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum Likelihood from Incomplete Data via the EM Algorithm," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 39, no. 1, pp. 1–38, 1977.
- [17] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
- [18] W. Xu, M. Zhang, A. Sawchuk, and M. Sarrafzadeh, "Co-recognition of human activity and sensor location via compressed sensing in wearable body sensor networks," in *Wearable and Implantable Body Sensor Networks (BSN), 2012 Ninth International Conference on*, May 2012, pp. 124 –129.
- [19] M. Zhang and A. A. Sawchuk, "A feature selection-based framework for human activity recognition using wearable multimodal sensors," in *Proceedings of the 6th International Conference on Body Area Networks*, ser. BodyNets '11, 2011, pp. 92–98.
- [20] W. Xu, M. Zhang, A. Sawchuk, and M. Sarrafzadeh, "Robust human activity and sensor location corecognition via sparse signal representation," *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 11, pp. 3169–3176, 2012.