2016 15th IEEE International Conference on Machine Learning and Applications

A Hierarchical Meta-Classifier for Human Activity Recognition

Anzah H. Niazi*, Delaram Yazdansepas†, Jennifer L. Gay‡, Frederick W. Maier*,
Lakshmish Ramaswamy†, Khaled Rasheed*†, Matthew P. Buman§
*Institute for Artificial Intelligence, The University of Georgia, Athens, Georgia
†Department of Computer Science, The University of Georgia, Athens, Georgia
‡College of Public Health, The University of Georgia, Athens, Georgia
§School of Nutrition and Health Promotion, Arizona State University, Phoenix, Arizona

***Abstract*—This paper proposes a multi-level meta-classifier for identifying human activities based on accelerometer data. The training data consists of 77 subjects performing a combination of 23 different activities and monitored using a single hip-worn triaxial accelerometer. Time and frequency based features were extracted from two-second windows of raw accelerometer data and a subset of the features, together with demographic information, was selected for classification. The activities were divided into five activity groups: non-ambulatory activities, walking, running, climbing upstairs, and climbing downstairs. Multiple classification techniques were tested for each classifier level and groups. Random forests were found to perform comparatively better at each level.**

Based upon those tests, a 3-level hierarchical classifier, consisting of 5 random forest classifiers, was built. At the first level, the non-ambulatory activities are separated from the rest. At the second, the ambulatory activities are divided into four activity groups. At the final level, the activities are classified individually.

Accuracy on test sets was found to be approximately 87% overall for individual activities and 94% at the activity group level. These results compare favorably to contemporary results in classifying human activity.

I. Introduction

Interest in human activity recognition has seen growth in recent years. Both the commercial market and the research sector have seen an increase in demand for efficient techniques in the classification of activities. The public health sector, in particular, is in need of such models. The target of this work is the development of a classification model for triaxial accelerometer data gathered from a demographically diverse group of subjects which could be used for test subjects under free-living conditions. Such a classification could be used to better estimate energy expenditure for individuals and determine the necessary steps for the individual to take towards a healthy lifestyle.

Activity recognition remains a difficult classification problem, mainly due to the sheer variance in how subjects perform the same activities and the relative difficulty in effectively monitoring subjects. Additional hardware (e.g., multiple body-worn accelerometers, heart-rate monitors, etc.) can improve classification, but this is typically not practical under free-living conditions. Using fewer measuring instruments is better from a subjects point of view, though it tends to be insufficient from a research perspective. A highly accurate classifier based on data from a single unobtrusive set of sensors would be ideal.

There are many ways to approach activity recognition. An important distinction to be made is between *within-subject* and *between-subject* recognition. For *within-subject* classification, individual models are produced for each subject. As is shown in Section 2, this can lead to very accurate results. However, by its very nature, the model development involves extensive training and calibration on each subject, which is difficult to achieve outside of controlled conditions.

In the case of *between-subject* classification, a universal model for all subjects is created. This model would (ideally) be trained only once and be applicable to all future subjects. Arguably, this is a more practical approach, as little calibration would be needed post-training. However, the generalized nature of the single model typically results in lower overall accuracy compared to individualized models.

For our research, the data used consists of triaxial data obtained from 77 subjects, quite fairly spread demographically, for 23 activities in studies performed at Arizona State University in 2013 and 2014. This is a fairly large data set compared to similar research in the field.

The ultimate objective of this research is the development of models to be used in free-living conditions by subjects for the purpose of health monitoring and energy expenditure calculation. This would eliminate the need for self-reporting as is customary in free-living studies which tends to be inaccurate due to the nature of human error. It should be noted that the data was collected in a controlled environment, though the subjects had significant choice in the selection and order of the activities they performed.

The classifier we have built uses data extracted from a single triaxial accelerometer, an ActiGraph GT3X+, along with demographic information from the subjects. Our approach was for between-subject classification and grouping similar activities into distinct groups. The final 3-level classifier was built using random forests to cater for each group of activities as well as the individual activities that they contain.

Section II provides analysis of prior research done in this domain. Section III explains how the data was collected and prepared for training. Section IV goes on to describe how the classifier was trained and the results obtained in testing. Finally, Section V considers potential areas for future research and the methods by which results could be further improved.

978-1-5090-6167-9/16 $31.00 © 2016 IEEE
DOI 10.1109/ICMLA.2016.16

TABLE I
DESCRIPTION OF ACTIVITIES PERFORMED

#	Activity	Duration or Distance	# of subjects
1	Treadmill at 27 młmin-1 (1mph) @ 0% grade	3 min	29
2	Treadmill at 54 młmin-1 (2mph) @ 0% grade	3 min	21
3	Treadmill at 80 młmin-1 (3mph) @ 0% grade	3 min	28
4	Treadmill at 80 młmin-1 (3mph) @ 5% grade (as tolerated)	3 min	29
5	Treadmill at 134 młmin-1 (5mph) @ 0% grade (as tolerated)	3 min	21
6	Treadmill at 170 młmin-1 (6mph) @ 0% grade (as tolerated)	3 min	34
7	Treadmill at 170 młmin-1 (6mph) @ 5% grade (as tolerated)	3 min	26
8	Seated, folding/stacking laundry	3 min	74
9	Standing/Fidgeting with hands while talking.	3 min	77
10	1 minute brushing teeth + 1 minute brushing hair	2 min	77
11	Driving a car	-	21
12	Hard surface walking w/sneakers	400m	76
13	Hard surface walking w/sneakers hand in front pocket	100m	33
14	Hard surface walking w/sneakers while carry 8 lb. object	100m	30
15	Hard surface walking w/sneakers holding cell phone	100m	24
16	Hard surface walking w/sneakers holding filled coffee cup	100m	26
17	Carpet w High heels or dress shoes	100m	70
18	Grass barefoot	134m	20
19	Uneven dirt w/sneakers	107m	23
20	Up hill 5% grade w high heels or dress shoes	58.5m x 2 times	27
21	Down hill 5% grade w high heels or dress shoes	58.5m x 2 times	26
22	Walking up stairs (5 floors)	5 floors x 2 times	77
23	Walking down stairs (5 floors)	5 floors$\times$2 times	77

II. RELATED WORK

Data mining and machine learning techniques have been applied extensively to physical activity classification. Commonly, classification is based on body-worn accelerometer data.

One of the initial problems in data mining is determining, for a given domain, the most suitable classifier type as well as the optimal features to use for classification. Preece et al. [11] provide a good overview of the tools that prove useful in activity classification as well the features that are frequently used. They present a comparison of different studies consisting of different routines, features, accelerometer placement locations and classification techniques. Some of the techniques analyzed include artificial neural networks, decision trees, k-nearest neighbor and support vector machines. A hierarchical system using thresholds is also discussed.

Bao & Intille [1] use five biaxial accelerometers for experiments, trained on 20 subjects and 20 activities in semi-natural settings. Decision trees gave them their best result of 84%. Additional experiments were carried out using only two accelerometers; it was found that a thigh- and hip-worn device combination provided the least decrease in accuracy.

In Ravi et al. [12], feature selection was used to identify that the mean, standard deviation, energy, and correlation are the best features for activity recognition. This study used data collected from two subjects wearing waist-mounted triaxial accelerometers. 8 activities were performed over the course of several days. Accuracy levels of 99% were obtained using Plurality Voting from a number of base classifiers trained by K-Nearest-Neighbors, Decision Trees, Naive Bayes, and Support Vector Machines. Though obtaining very high results, the study was only carried out on 2 subjects, which would not translate well to a general population.

Lester et al. [8] describe a personal activity recognition system using a custom-made module worn on the shoulder, consisting of an accelerometer, microphones and barometers (a combination of 7 devices). They combined multiple static classifiers using a Hidden Markov Model (HMM) and claim an accuracy of 90% (though their test accuracy reaches a high point of 84%). They tested their model on 12 subjects with 8 activities and indicate that accuracy drops to around 65% if only the accelerometer is used.

Yang et al. [15] use neural classifiers to classify eight domestic activities with data gathered from 7 subjects by wrist-worn accelerometers. They achieved a *within-subject* average accuracy of 95.24% by initially separating dynamic activities (running and walking) from static activities (standing and sitting) before using separate feature subsets for both types of activities.

Staudenmayer et al. [13] used artificial neural networks to classify groups of activities. They gather multiple activities into five groups (low level/non-ambulatory, locomotion, household activities, and vigorous sports). They had 48 subjects equally divided across genders and used triaxial accelerometers mounted at the waist. The accuracy of this system was 88.8% at the activity group level. The activities were not classified at the individual level.

Khan et al. [5] classifies a group of 15 activities consisting of 3 static (non-ambulatory) activities and 3 dynamic (ambulatory) activities with the rest being transitional activities that do not relate to this study. The data was collected from an accelerometer set at 20Hz worn on the chest by 6 subjects performing a specific sequence of activities each day for a month. They used a somewhat hierarchical approach to initially distinguish between static and dynamic activities followed by utilizing artificial neural networks and an augmented-feature vector to achieve accuracies averaging 97.9%. The data was collected on relatively few subjects for few activities with much attention given to transitional activities which are not the focus of our study. Additionally, the data was collected in specific activity sequences rather than the freer method applied by the subjects in our dataset. However, the paper displays the improvement in accuracy achieved by a hierarchical approach by comparing results obtained at a single-level classification (71.6%).

Kwapisz, Weiss & Moore [7] use the data mining software WEKA on a data set of 29 subjects performing 6 activities. The data collected came from a pocketed Android phone application. The activities performed would be regarded as high-level (grouped) activities in our work. They achieved an average accuracy of 91.7% using a multilayer perceptron.

Weiss et al. [14] developed a smartphone-based system using random forests on 5 grouped activities walking, jogging, climbing stairs, Standing, and sitting/lying down. They use both personalized (*within-subject*) and universal (*between-subject*) models. They extracted 43 features from triaxial accelerometer data. The universal model was formed in Lockhart & Weiss [9] and shows an accuracy of 76% with similar results on new subjects. The personalized data is said to have accuracies generally higher than 95% for new subjects who train on themselves.

Phan [10] uses a pruned decision tree to classify 5 (grouped) activities performed by 20 subjects. The data was gathered using a Samsung mobile phone at 32 Hz. They achieved an accuracy of 96.8% by pruning off and discarding data after training with a C4.5 decision tree.

Deng et al [2] uses various KELMs (kernel extreme learning machines) on a dataset of 30 subjects performing 6 activities collected at 50Hz on a waist-worn smartphone accelerometer. They achieved accuracies of 99% using this approach.

Zheng [17] features a hierarchical approach similar to ours in classifying HAR. The dataset consists of hipworn accelerometer data collected at 100Hz from 14 subjects on 10 self-paced activities. The study divides the activites into 4 "states". Using multilayered Least Squares Support Vector Machines (LS-SVM) and Naive Bayes (NB) classifiers, they obtained an average accuracy of 95.6% across individual activities.

Our study, compared to those just discussed, consists of a significantly larger dataset with activities categorized at a more sophisticated level than most datasets implemented by other studies. The primary reason for this is that this study is aimed at recognizing activities for public health purposes rather than general classification. While this certainly provides a bigger challenge at achieving high accuracies on this dataset, we intend to show that a hierarchical approach significantly improves our results over a single-layer classification and compares favorably to other studies despite the higher level of discrepancy involved in activity separation and the use of a single accelerometer.

III. Data Collection, Preprocessing and Experiments

Our data was trained on 77 subjects performing 4 grouped and 23 individual activities. This section details the collection and preprocessing of this data.

A. Participants and procedures

Participants were recruited from the Phoenix, AZ and surrounding areas through community sources, email distribution lists, and social media outlets. Participants were 18-64 years of age and free of any contraindications for exercise. Participants were fitted with the accelerometer and completed a series of activities for 3 min in duration (see Table I). Virtually all participants completed the following activities: standing, fidgeting with hands while talking; 1 min of brushing teeth and 1 min brushing hair; some form of hard surface or carpet walking; and walking up and down stairs. An additional three treadmill activities and three other activities were randomly assigned. Timestamps for the beginning and end of activities were captured using a custom-built Android application which was synced to the same computer as the activity monitor.

B. Activity monitoring

Participants were fitted with the ActiGraph GT3X+ (ActiGraph LLC, Pensacola, FL) activity monitor positioned along the anterior axillary line of the non-dominant hip. The monitor was fixed using an elastic belt. The ActiGraph GT3X+ is a lightweight monitor (4.6cm x 3.3cm x 1.5 cm, 19g) that measures triaxial acceleration ranging from -6g to +6g. Devices were initialized to sample at a rate of 100hz. Accelerometer data were download to and extracted using Actilife 5.0 software (ActiGraph, LLC, Pensacola, FL).

310 subjects participated in the study. From them, data from 77 subjects, 55 females and 22 males, were used to train our classifiers. Table II provides demographic information on the subjects.

TABLE II
SUBJECT DEMOGRAPHICS

	Mean	Standard Deviation	Range
Age (Years)	33.2	9.7	18.2 - 63.2
Height (cm)	167.9	7.9	152.6 -188.9
Weight (kg)	72.1	12.1	48.3 - 105.5
BMI	25.6	3.9	17.7 - 35.4

C. Feature Extraction

A total of 246 features were initially extracted from the raw data. A summary of the features is as follows:

- **Features in the time domain**: features extracted from the axes and their first differentials and the vector magnitude. These include the mean, maximum, minimum values, standard deviations, median crossings and the 10th, 25th, 50th, 75th, 90th percentiles. Also included are the correlations between the each axes as well as the correlations between their first differentials.

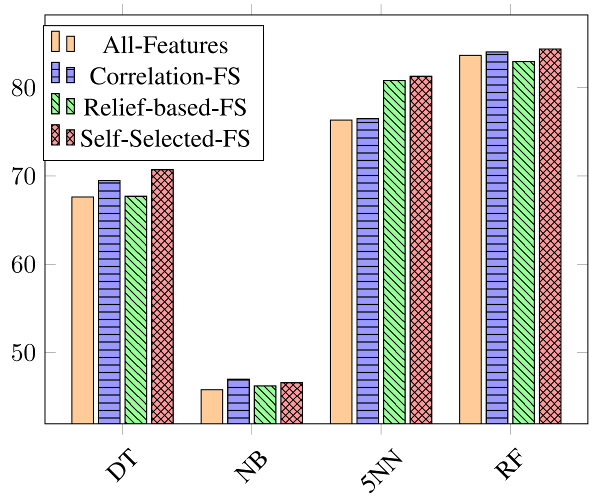

Fig. I. Accuracy performance for feature subsets across four classifiers.DT = Decision Trees; NB = Naive Bayes; 5NN = 5 Nearest Neighbor; RF = Random Forests,100 trees

- **Features in the frequency domain**: these include the dominant frequency and its magnitude for the axes, their first differentials and the vector magnitude.
- **Features of wavelet analysis**: features extracted from the 1st level to the 5th level of wavelet decomposition coefcients for each accelerometer signal (x,y and z). These include the mean, maximum, minimum, standard devi-ations, median crossings and the 10th, 25th, 50th, 75th, 90th percentiles.
- **Demographic features**: these include the age, height, weight, gender and BMI of the subject.

D. Feature Selection

The initial 246 features, while helpful for initially gauging the data, increase complexity. Generating the features requires extra work and significantly increases the time needed to train and run a classifier. Feature selection was carried out to whittle down the features to a smaller but more significant subset.

Initially, correlation-based [3] and relief-based [6] feature selection methods were used. Additionally, a 42-feature subset was selected using domain knowledge and an eye to standardize the preprocessing step. This feature set performed comparably well to the others in preliminary testing and was subsequently the feature set used in creating the final classifier. Figure I shows the classifiers performance on the feature subsets compared to the entire feature set.

Random Forests (100 trees) trained on the self-selected subset outperformed the other subsets across the classifiers. While other subsets also performed substantially well with Random Forests, the self-selected was also a robust performer with other classifiers, signifying it's strength as a subset.

E. Activity Groups

The 23 activities were separated into 5 activity groups, each group containing similar types of activities. Grouping permits layered classification—a higher level classifier learns to efficiently distinguish between groups of activities, while lower level classifiers specialize in classifying in-group activities. A possible disadvantage of this arrangement is that misclassification of instances at higher levels can propagate errors to lower levels. Because of this, it is imperative that the higher-level classifiers be much more accurate.

The activities were grouped as shown in Table III. Activities which involved minimal physical activity were classified as non-ambulatory. Ambulatory activities were further divided into 4 subgroups; walking (all locomotive & treadmill activities less than 4 mph), running (treadmill activities more than 4 mph), climbing upstairs, and climbing downstairs.

TABLE III
DIVISION OF ACTIVITIES IN THE CLUSTERS

Non-Ambulatory Activities	
8,9,10,11	
Ambulatory Activities	
Walking	1,2,3,4,12,13,14,15, 16,17,18,19,20,21
Running	5,6,7
Upstairs	22
Downstairs	23

IV. EXPERIMENTAL RESULTS

WEKA [3], a data mining software application, was used for classifier development. WEKA provides extensive libraries of machine learning techniques and a easily manageable GUI to carry out training experiments.

Though random forests was indicated as a preferable base classifier during the feature selection, other preliminary tests were carried out. A 16 subject subset was used to train many different classifiers with 10-fold cross-validation (as shown in Figure II). In these initial tests, a 100 tree Random Forest outperformed all other classifiers, including meta-classifiers such as bagged decision trees and stacked classifiers.

Random Forests [4] work by generating many decision trees and classifying according to the modal class of the "forest". Subsequent experiments at the activity group level repeatedly indicated that random forests were a reasonable choice for at all levels of the meta-classifier.

Our data set contained 87,943 records. 90% of the data was used as training data and the remaining 10% was reserved as the test set, the data separated by stratification. Due to high number of activities in the walking group, the training set was weight-balanced to avoid a bias.

Five random forest classifiers were obtained for level 1, 2 and 3 for the walking, running and non-ambulatory groups. The trained classifiers were then set up in a Java program. The test set was run through this Java program which resulted in confusion matrices for each level. The accuracy, precision and recall for each level and activity were extracted from the confusion matrices using standard equations.

$$Precision = \frac{tp}{tp + fp} \tag{1}$$

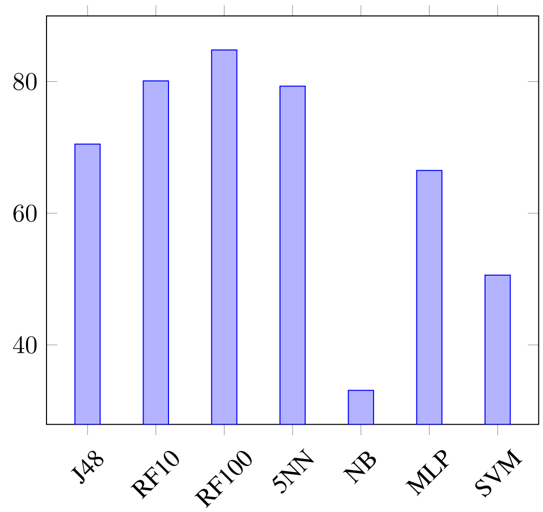

Fig. II. Performance of classification techniques on the 16 subject training set with 10-fold cross validation. J48 = C4.5 decison tree; RF10 = Random Forest, 10 trees; RF100 = Random Forest 100 trees; 5NN = 5 Nearest Neighbor; NB = Naive Bayes; MLP = Multilayer perceptron; SVM = Support Vector Machine

$$Recall = \frac{tp}{tp + fn} \tag{2}$$

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \tag{3}$$

$$F - measure = 2\frac{precision \times recall}{precision + recall} \tag{4}$$

tp=true positive, fp=false positive, fp=true negative, fn=true negative.

A. LEVEL ONE

At level one, the activities are divided into ambulatory and non-ambulatory activities. The accuracy at this level is 97.899%. The misclassified instances will contribute to some trickle-down errors in the levels below. However, the accuracy is low enough for overfitting not to be an issue. As the data set does not account for transitional activities (e.g., transitioning from standing to running), these discrepancies are welcome. The performance of level one is shown in table IV.

TABLE IV
PERFORMANCE ANALYSIS FOR LEVEL 1

LEVEL 1	Precision	Recall	F-Measure
Non-Ambulatory	0.987	0.944	0.965
Ambulatory	0.978	0.995	0.986
Accuracy: 98.03%			

B. LEVEL TWO

At level two, the activities are divided into groups, shown in Table V. Note that the non-ambulatory activities were separated at level one. Therefore, the errors shown for that group are a result of the trickle-down effect of level one. The accuracy at level 2 is almost 94%, which we believe compares favorably to other research, e.g., the 88% of Staudenmayer, et al. [13] and the 91.7% of Kwapisz, Weiss & Moore [7]. Phan [10], Ravi et al. [12], Khan et al [5] and [17] all achieve higher accuracies on their datasets, but these studies involve fewer subjects and overall less data. Given these discrepancies, our results are certainly competitive.

The results in Table V can also be compared to results based on models using multiple sensors. E.g., Lester et al. [8] achieved 84% with 7 sensors. Bao & Intille [1] achieved 84% as well with 5 sensors.

TABLE V
PERFORMANCE ANALYSIS FOR LEVEL 2

LEVEL 2	Precision	Recall	F-Measure
Non-Ambulatory	0.944	0.987	0.965
Walking	0.915	0.978	0.945
Running	0.993	0.847	0.914
Upstairs	0.958	0.813	0.880
Downstairs	0.969	0.755	0.849
Accuracy: 93.673%			

C. LEVEL THREE

At this level, activities are classified individually inside their respective groups. Overall, the accuracy obtained was 86.63%. Generally, other classification studies do not classify at this level due to similarity and overlap between activities. By activity grouping, we can train classifiers specifically for a group of activities which would otherwise be difficult to separate. The results for the non-ambulatory, walking and running groups are shown in Tables VI, VII and VIII.

TABLE VI
PERFORMANCE ON NON-AMBULATORY ACTIVITIES

Non-Ambulatory Activities	Precision	Recall	F-Measure
Seated, folding/stacking laundry	0.924	0.951	0.937
Standing/Fidgeting while talking	0.933	0.958	0.945
brushing, 1min teeth + 1min hair	0.934	0.851	0.891
Driving a car	0.990	0.997	0.993
Overall Accuracy: 93.739%			

TABLE VII
PERFORMANCE ON RUNNING ACTIVITIES

Running Activities	Precision	Recall	F-Measure
Treadmill 5mph @ 0% grade	0.963	0.991	0.977
Treadmill 6mph @ 0% grade	0.976	0.965	0.970
Treadmill 6mph @ 5% grade	0.979	0.971	0.975
Overall Accuracy: 97.368%			

Note that as the "upstairs" and "downstairs" groups are single-activity groups, table results for them are not shown. It should be noted that the "upstairs" activity group had an accuracy of 81.33% and the "downstairs" activity group had an accuracy 75.5%. The downstairs activity has significantly lower recall, perhaps due to the varied approach subjects would have to climbing downstairs, demographically. For example, older and/or weightier subjects would descend stairs much more slowly than younger and/or lighter subjects.

As no literature has been found that classifies activity at such a minute level, we do not have a precedent to compare our

TABLE VIII
PERFORMANCE ON WALKING ACTIVITIES

Walking Activities	Precision	Recall	F-Measure
Treadmill 1mph @ 0%	0.900	0.997	0.946
Treadmill 2mph @ 0%	0.904	0.943	0.923
Treadmill 3mph @ 0%	0.917	0.871	0.893
Treadmill 3mph @ 5%	0.932	0.927	0.929
Hard surface	0.842	0.967	0.900
Hard surface, hand in pocket	0.956	0.752	0.842
Hard surface, carrying 8 lbs.	0.913	0.652	0.761
Hard surface, cell phone	0.921	0.648	0.761
Hard surface, coffee	0.852	0.730	0.786
Carpet, heels/dress shoes	0.869	0.816	0.842
Grass barefoot	0.944	0.878	0.910
Uneven dirt w/sneakers	0.957	0.611	0.746
Uphill 5%, heels/dress shoes	0.948	0.895	0.921
Downhill 5%, heels/dress shoes	0.946	0.859	0.900
Overall Accuracy: 88.722%			

results to. However, section V will elaborate on other studies being carried out on this dataset

Non-Ambulatory Activities: Non-ambulatory activities are those that require little to no locomotion. The hierarchical approach allows the individual classifier to exemplify the minute differences between activities with a relatively low distinguishing features, achieving an accuracy of 94%. Table VI shows the performance of non-ambulatory activity classification.

Running Activities: Treadmill activities at 5-6 mph were regarded as running activities. The classifier achieved an very high accuracy of 97%, shown in table VII.

Walking Activities: The overall accuracy for classification of walking activities is 88.7%. Table VIII shows the comparatively low recall rates of the "Hard surface" activities, indicating that these activities are much more difficult to separate. The difference between these activities is arm position (holding cell phone, coffee cup, etc.) a feature that would probably be better detected by a wrist-worn accelerometer than a hip-worn one.

V. CONCLUSION AND FUTURE WORK

The hierarchical meta-classifier achieved an accuracy of 93.7% at the activity group level, which compares favorably to other group-level studies discussed in section II. Furthermore, the meta-classifier was able to distinguish between intra-group activities with an accuracy of 86.6% as compared to 84% achieved with Random Forests at a single level (Figure II).

Future work includes testing and calibrating the classifier on free-living data to calculate energy expenditure information for health informatic purposes. Further investigation involves the effect of sampling rates and window sizes on classification accuracy, a comparative study [16] which dives into a demographic and classifier based analysis, and classifying the data set using deep neural network, as well as other possible approaches.

REFERENCES

[1] Ling Bao and Stephen S. Intille. *Pervasive Computing: Second International Conference, PERVASIVE 2004, Linz/Vienna, Austria, April 21-23, 2004. Proceedings*, chapter Activity Recognition from User-Annotated Acceleration Data, pages 1–17. Springer Berlin Heidelberg, Berlin, Heidelberg, 2004.

[2] Wan-Yu Deng, Qing-Hua Zheng, and Zhong-Min Wang. Cross-person activity recognition using reduced kernel extreme learning machine. *Neural Networks*, 53:1–7, 2014.

[3] Mark A. Hall. Correlation-based feature selection for discrete and numeric class machine learning. In *Proceedings of the Seventeenth International Conference on Machine Learning*, ICML '00, pages 359–366, San Francisco, CA, USA, 2000. Morgan Kaufmann Publishers Inc.

[4] Tin Kam Ho. Random decision forests. In *Document Analysis and Recognition, 1995., Proceedings of the Third International Conference on*, volume 1, pages 278–282, Aug 1995.

[5] Adil Mehmood Khan, Young-Koo Lee, Sungyoung Y. Lee, and Tae-Seong Kim. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer. *IEEE Transactions on Information Technology in Biomedicine*, 14(5):1166–1172, 2010.

[6] Kenji Kira and Larry A. Rendell. A practical approach to feature selection. In *Proceedings of the Ninth International Workshop on Machine Learning*, ML92, pages 249–256, San Francisco, CA, USA, 1992. Morgan Kaufmann Publishers Inc.

[7] Jennifer R. Kwapisz, Gary M. Weiss, and Samuel A. Moore. Activity recognition using cell phone accelerometers. *SIGKDD Explor. Newsl.*, 12(2):74–82, March 2011.

[8] Jonathan Lester, Tanzeem Choudhury, and Gaetano" Borriello. Pervasive computing: 4th international conference, pervasive 2006, dublin, ireland, may 7-10, 2006. proceedings. In *A Practical Approach to Recognizing Physical Activities*, pages 1–16. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006.

[9] Jeffrey W. Lockhart and Gary M. Weiss. The benefits of personalized smartphone-based activity recognition models. In *Proceedings of the 2014 SIAM International Conference on Data Mining*, pages 614–622. 2014.

[10] Thomas Phan. Improving activity recognition via automatic decision tree pruning. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*, UbiComp '14 Adjunct, pages 827–832, New York, NY, USA, 2014. ACM.

[11] Stephen J Preece, John Y Goulermas, Laurence P J Kenney, Dave Howard, Kenneth Meijer, and Robin Crompton. Activity identification using body-mounted sensorsa review of classification techniques. *Physiological Measurement*, 30(4):R1, 2009.

[12] Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman. Activity recognition from accelerometer data. In *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3*, IAAI'05, pages 1541–1546. AAAI Press, 2005.

[13] John Staudenmayer, David Pober, Scott Crouter, David Bassett, and Patty Freedson. An artificial neural network to estimate physical activity energy expenditure and identify physical activity type from an accelerometer. *Journal of Applied Physiology*, 107(4):1300–1307, 2009.

[14] Gary M. Weiss, Jeffrey W. Lockhart, Tony T. Pulickal, Paul T. Mchugh, Isaac H. Ronan, and Jessica L. Timko. Actitracker: A smartphone-based activity recognition system for improving health and well-being, 2014.

[15] Jhun-Ying Yang, Jeen-Shing Wang, and Chen Yen-Ping. Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers. *Pattern Recognition Letters*, 29(16):2213–2220, 2008.

[16] Delaram Yazdansepas, Anzah H. Niazi, Jennifer L. Gay, Frederick W. Maier, Lakshmish Ramaswamy, Khaled Rasheed, and Matthew P. Buman. Towards robust human activity recognition systems; a wearable sensor-based multi featured approach, 2016.

[17] Yuhuang Zheng. Human activity recognition based on hierarchical feature selection and classification framework. *Journal of Electrical and Computer Engineering 2015*, pages 34–43, 2015.