2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE)

Smartwatch Based Activity Recognition Using Active Learning

Farhad Shahmohammadi, MS; Anahita Hosseini; Christine E. King, PhD; Majid Sarrafzadeh, PhD
Department of Computer Science
University of California, Los Angeles
Los Angeles, CA 90095
{farhad, anahosseini, kingce, majid} @cs.ucla.edu

***Abstract*—Human activity monitoring has become widely popular in recent years, and has been utilized in a vast number of fields and applications. Most of the activity recognition algorithms proposed have emphasized the use of inertial sensors in smartphone devices or other bodily-worn sensors. However, wearable inertial sensors are not interactive, and smartphones are not easily worn. Thus, with the advancement of smartwatches, unique opportunities exist to provide user interaction and highly accurate personalized activity recognition. Through the use of Active Learning, an interactive machine learning technique, specific behaviors can be learned by querying for unknown actions. This paper describes a smartwatch-based active learning method for activity recognition to identify 5 commonly performed daily activities. The results of this study revealed that this system can obtain a 93.3% accuracy across 12 participants. From our results, we demonstrate that an interactive learning approach using active learning in smartwatches has significant advantages over smartphones and other devices for activity recognition tasks.**

Keywords—activity recognition; smartwatch; active learning; inertial sensors; wearable devices

I. INTRODUCTION

Human activity monitoring has become widely utilized in recent years among a vast number of fields [1, 2], particularly in healthcare [3, 4, 5, 6, 7]. As [8] explains, recognizing and monitoring activities such as walking and running in patients with chronic disease conditions such as diabetes, obesity or heart disease is necessary to monitor treatments and improvements in behaviors. Moreover, as [9] suggests, prevention of many health disorders in children is possible by monitoring their activity patterns. Currently, most smartphones and smartwatches are equipped with motion and direction sensors, which can be utilized to identify human activities. However, most effort has focused on smartphone based activity recognition, such as those presented in [10, 11, 12, 13].

There have been few studies that have focused on smartwatches. For instance, [14] used a combination of smartwatch and smartphone for improving activity recognition accuracy. The lack of attention towards these devices may be due to several reasons. First, it hasn't been until recent years that smartwatches have become popular among the general public [15]. Also, Android-based smartwatches have been limited in terms of computation power, battery life, and have lacked WiFi support prior to the Android Wear 5.1.1 update. Prior to this update, smartwatches were only able to connect to the Internet through the smartphone, requiring them to be proximally tethered. Thus, in the absence of a smartphone or device with WiFi or other form internet connectivity, smartwatches were not able to play a gateway role by performing data collection and sending this information to a web server to perform computations in real time. As a result, it was previously impractical to use smartwatches as the sole device in large scale and complex activity recognition applications. However, since many of these issues are now resolved and smartwatches are becoming more popular, they have become a promising tool for activity recognition applications.

Smartwatches pose several advantages over other wearable inertial sensors and smartphones. First, smartwatches combine features of smartphones with continuous data monitoring [16]. By providing a screen similar to those seen in smartphones, smartwatches can provide interactive feedback to allow for direct communication with the user from any location [17]. They are also ubiquitous in that they are typically worn even when at home and during nighttime. Furthermore, unlike smartphones that are bulky and not always worn by the individual during behaviors of interest such as exercise, smartwatches can easily be worn during high levels of activity to provide continuous sensing information beyond accelerometry, such as heart rate, global positioning satellite (GPS), gyroscope, and compass data [18]. This is particularly promising in applications that require continuous activity monitoring to identify unexpected changes in behavior patterns and propose alarms and guidance based on the given localized area. The messages and alarms delivered to the user are also more easily observed than those sent to smartphones, as individuals can receive vibrations, text, and sounds within immediate proximity to their line of sight. Finally, the modularity of apps that can be deployed on smartwatches provides an unlimited resource regarding their use in physical activity monitoring and other applications.

To provide an accurate estimate of a wide range of activities from smartwatches, active machine learning algorithms can be utilized to provide improvements to the activity recognition model. Active Learning provides queries to the individual to annotate data points and thus provide valuable information to improve the current model. In this context, it requires individuals to select the activity that was being performed during data points that are unknown, thus increasing the number of informative samples and provide an individualized training model for various physical activities [19].

In this study, we demonstrate that smartwatches can identify daily activities, and can become an invaluable tool for future

978-1-5090-4722-2/17 $31.00 © 2017 IEEE
DOI 10.1109/CHASE.2017.115

activity recognition applications. Furthermore, we show how active machine learning models can be leveraged in smartwatches to provide personalized models of each individual and improve the accuracy of the activity recognition classifier. Since smartwatches provide an easy to read interface to interact with individuals, active learning methods were easily deployed on this device to provide individualized and continuous activity recognition monitoring. We also propose a strategy for the balancing number of queries that is made through the smartwatch. Finally, we show that by employing only two Android-based sensors, we can improve the classification accuracy compared to commonly used raw inertial sensors.

II. Related Work

A. Activity Recognition

Human activity recognition from inertial sensors has been widely studied. Historically, initial work in this area such as those described in [20, 21, 22, 23] monitored activity using multiple inertial sensors that were mounted on different parts of the body. However, researchers found that this was not cost effective nor user friendly [24], and with the advancement of smartphone technology, activity recognition studies began to rely on the use of smartphones to measure activities from inertial sensors. For example, [10, 11] used smartphones and basic machine learning models such as Random Forests and Support Vector Machine to classify activities of users. In addition, [25] reviewed and compared different supervised algorithms, placement of inertial sensors, and their impact on activity recognition accuracy. These methods, although provided accurate results on the collected dataset, were not personalized and required a large dataset to classify different behaviors. However, as [3] explains, in the healthcare domain, we are confronted with small datasets and rare events. Furthermore, streams of data encounter temporal changes and this cannot be modeled using a fixed dataset [19].

To solve aforementioned problems, some studies such as the study presented in [26] used semi-supervised learning to reduce the size of the dataset. However, semi-supervised learning methods do not allow for personalized models for individuals who exhibit varying behaviors, as their dataset is not updated during training. Active Learning gives us the opportunity to train the model with each user's data on selected and limited training samples.

B. Active Learning

Unlike more commonly used supervised machine learning methods, Active Learning allows us to generate activity recognition models with significantly smaller training datasets. This is particularly beneficial in applications that target the elderly, children, and those with disease conditions, as performing certain activity tasks such as exercise for long periods of time may be too difficult in these populations. The key hypothesis in Active Learning is how to choose training instances and ask for their label such that the model learns improves its accuracy with fewer training samples.

The first family of active learning methods that was studied was Membership Query Synthesis [27]. As [28] explains, this approach generates arbitrary combinations of features and queries for labels. However, because many of these samples are not collected from real world behaviors, collecting labels from human user is difficult and awkward to perform.

Query by Committee [29] methods are a more recently studied approach that have been used widely in real-world applications. In this method, a committee of models is created to vote for the label of each sample. The sample with the largest disagreement is considered the most informative sample, and thus used to query the individual for a label.

Other Active Learning methods mainly rely on different query strategies that decide when to make a query for label of an instance. For instance, one the most popular query strategies that has several variations and has been widely used is Uncertainty Sampling [30]. In this method, the uncertainty of the label is defined as a measure and those samples with the highest uncertainty are selected to query the user for labels.

Many of the initial applications of Active Learning have been pool-based. In this approach, all training samples are available prior to testing and the most informative samples are selected to be queried for a label. However, in many real-world applications, such as in activity recognition problems, stream-based Active Learning is required. This is because users cannot remember their activity history for a long period of time and cannot determine the activity that was performed simply by looking at the recorded sensor data at the end of the recording period. As a result, stream-based active learning must be utilized to decide whether to query for an activity label upon the arrival of each sample in real time, and then improve its model instantly. This method requires instant and easy user interaction capabilities in the activity recognition system.

Few studies have focused on the stream-based nature of inertial sensor data for activity recognition problems. In [31], a smartphone activity recognition system was proposed that employed a stream-based segmentation technique. However, their active learning model was reliant on hyper-parameters that required a large annotated dataset to be calculated, which is counter intuitive in Active Learning. Another smartphone based activity recognition framework was introduced in [19] in which a lightweight clustering-based active learning model was used to predict activity. However, their model did not place any limitations on the number of queries sent to the user. This is problematic, as [32] suggests, as a high number of queries will reduce compliance with use of the system. Thus, an activity dependent query strategy must be utilized to minimize unnecessary queries and reduce user burden.

To the best of author's knowledge, no study has used active learning on smartwatches for activity recognition problems. However, with advancement of smartwatches, few studies such as [15] have shown that new generations of smartwatches can replace smartphones for activity recognition problems, as interaction capabilities of smartwatches for improving activity recognition has previously been neglected. These newer smartwatches are promising in that they can provide queries to users performing activities for Active Learning methods, thus providing highly accurate and personalized activity recognition models.

III. Background

Prior to describing our methods, we will first introduce some basic concepts that will be used throughout the study.

A. Software-based Sensors

As described in reviews such as the one described in [33], most supervised activity recognition models use raw hardware-based inertial sensor features. However, the Android framework provides a set of software-based sensors that to the best of authors knowledge, has yet to be assessed in activity recognition problems. These sensors have also been used in [34] for recognizing transportation modes, which had promising results. In addition, they provide simpler features, which aids in reducing the power consumption of the activity recognition algorithm implemented on the smartwatch. In the following sections, we will briefly introduce these commonly used sensors in Active Learning.

1) Linear Acceleration Sensor

The linear acceleration sensor provided by Android smartwatches utilizes a tri-axial accelerometer that measures the acceleration applied to it in the X, Y, and Z-axes relative to the device's coordinate system (Fig. 1). The acceleration applied to the smartwatch that is read by the raw accelerometer sensor includes the force of gravity, which is used to determine the orientation of the smartwatch on the individual's wrist. However, acceleration due to gravity makes it difficult to find the acceleration solely due to the user's activity. As a result, the linear acceleration sensor excludes the force of gravity from the accelerometer's measurement and reports only the acceleration applied to each axis by the individual.

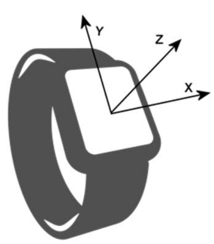

Fig. 1 Embedded accelerometer and its axes in a smartwatch.

2) Rotation Vector Sensor

The rotation vector sensor reported by Android is a fused sensor that measures the orientation of the device relative to the East-North-Up coordinate system. This vector sensor mainly uses integration over the gyroscope on the smartwatch to obtain the orientation of the smartwatch. In addition, the sensor incorporates information from the watch's accelerometer and magnetometer sensors to obtain a more accurate estimation of the current orientation. The rotation vector sensor then returns the smartwatch's orientation as a unit quaternion. The quaternion is a four-dimensional extension of the complex numbers system, which is very useful for spatial rotations [35]. Specifically, a rotation of angle θ around the obtained axis will transform the East-North-Up coordinate system to the smartwatch's relative coordinate system using the form:

$$(\cos(\theta/2),\ x\cdot\sin(\theta/2),\ y\cdot\sin(\theta/2),\ z\cdot\sin(\theta/2)) \tag{1}$$

where (x, y, z) in (1) is the axis of rotation in three dimensions.

B. Active Learning

Active Learning is a learning strategy that learns based on query and the main idea is to query only when model can learn something new and improve performance. There are two main advantages in this approach, as described in [15]. First, for many systems, labeling instances is very difficult and time-consuming, and active learning can eliminate the need for excessive labeling. Second, for applications like activity recognition, it allows for better personalization as the system learns based on subjects' answers to queries. Specifically, Active Learning algorithms define a strategy to decide when to query for label of a sample and use it as a training sample. Here, we explain two commonly used querying approaches that were also utilized in our study.

1) Uncertainty Sampling

The uncertainty sampling query strategy [36] allows the active learner to issue queries for those samples that model is least certain about their labels. Uncertainty can be defined based on application. For instance, as [15] exemplifies, in binary classification, one can define uncertainty as how near the posterior probability of being positive is to 0.5 and then query for points with uncertainty less than a threshold. In multi-class classification, after calculating uncertainty measure for each label, it is required to define a query strategy based on these uncertainties. For instance, one can query for the instance that has the least confidence as being the most probable label. It should be noted that this algorithm can be used both as a pool-based and stream-based method for defining the uncertainty measure.

2) Query by Committee

In the Query by Committee strategy [37], a committee of models is defined and each member votes for whether to issue a query. In this approach, the most informative data point is the one that has the most disagreement among its votes. These committee classifiers and voting system can be designed based on the desired application. Furthermore, Query by Committee is a pool-based method as it compares all data points at each step. Thus, modification is required to use it in a stream-based Active Learning model.

In this study, we employ a variation of these two commonly used Active Learning algorithms and query strategies in a stream-based activity monitoring application. Through the use of smartwatches, queries were given to the user in real time given two query strategies that minimize user burden. The remaining sections will describe how this was accomplished and tested among several participants wearing smartwatches.

IV. METHODS

We developed an Android-based smartwatch application that can be installed on any smartwatch to record and determine several commonly performed activities from individuals in real time. A sample of the application's user interface is depicted in Fig. 2. To test this application, we used the Samsung Gear Live (Samsung Electronics, Samsung, Seoul, South Korea) smartwatch that streamed inertial data at a frequency of 10 Hz to a webserver for online analysis. First, the performance of the Android software-based sensors was examined (the linear acceleration and rotation vector sensors), followed by exploring various Active Learning models through a small pilot study for activity recognition.

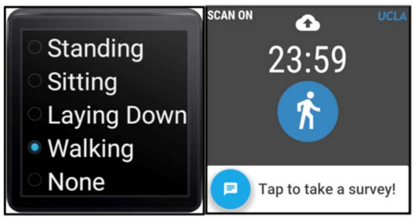

Fig. 2 Smartwatch user interface used to collect activity data by the participants.

A. Experiment Setup

Twelve individuals participated in the study, consisting of 8 males and 4 females between the ages of 22-28 years old. To collect activity data, each individual was asked via the smartwatch to select and perform one of five activities: running, walking, standing, sitting, and lying down. Note that the "None" activity option was designed so that the user could inform the application that an activity that is not listed was being performed such as cycling. Each activity was then performed for a total of 10 minutes, except for running which was performed for a total of 5 minutes to minimize discomfort and remove unwanted noise in the data when the participant was fatigued. The individuals performed each of these activities without the experimenter's supervision and provided a brief description of the actions performed during the activity. For example, during the standing activity, some users reported drinking and speaking while standing. In the sitting activity, individuals were allowed to move their upper body and some reported playing the piano or using a smartphone. During the lying down activity, some users also reported using a smartphone.

To evaluate the performance of the active learning model, the same activity dataset was used; however, the data was fed in a stream-based format into the model to simulate real-time activity recognition. On arrival of each data point, the active learner decided whether to query the point for a class label, and after each query, updated the model. To evaluate the model after each modification, each subject's data was divided into two equal length parts. Let these parts be described as sets A and B. In this phase, we treated set A as a time period in which the algorithm can issue queries. We then simulated the answers from the subject using our knowledge of the true labels for set A. Finally, we tested the updated active learning model on dataset B, which is further described below. To compare the accuracy of classification among different models, we used the F1 score, which is defined as:

$$F1 = 2 * \frac{Precision*Recall}{Precission+Recall} \qquad (2)$$

B. Feature Extraction

To extract important features for classification of activities, we first performed time-based segmentation over other techniques described in [38]. As [9, 39] suggests, 10 second window-sizes were chosen as this window size is a reasonable choice for activity recognition in young adults using smartphones with similar inertial sensors. For each window, a set of features was then generated that summarized the readings in this time window.

Another important feature to determine the activity being performed is the magnitude of the acceleration vector. This was calculated from the accelerometer readings and was generated as a new time series data. A summary of this data alongside the calculated rotation vector sensor, whose methods for calculation are described below, was also generated for each dataset.

In order to use the data collected from the rotation vector sensor, we transformed quaternions into an informative feature. To accomplish this, for each sample in each window, we rotated the (0,0,1) vector (the unit vector towards the sky) by each quaternion to obtain the (x,y,z) coordinates of the unit normal vector of the smartwatch surface. Although the unit normal vector of the smartwatch surface does not uniquely specify the orientation of the watch, its combination with other sensor data was able to infer the correct orientation. Furthermore, we noticed that including other vectors increased the complexity and worsened performance of the model, so these features were removed from the activity classification model. A list of resulting features that were used to generate summaries of the 10 second windows is provided in Table 1.

TABLE I. LIST OF FEATURES USED IN THE ACTIVITY CLASSIFIER.

Features	*Sensors*
Mean	Acceleration & Orientation
Standard Deviation	Acceleration & Orientation
Skewness	Acceleration & Orientation
Kurtosis	Acceleration & Orientation
Dynamic Time Warping Distance	Acceleration & Orientation
Energy	Acceleration
Inter Quatile Range	Acceleration
Average of the absolute differences between successive data points	Acceleration
Standard deviation of the absolute differences between successive data points	Acceleration

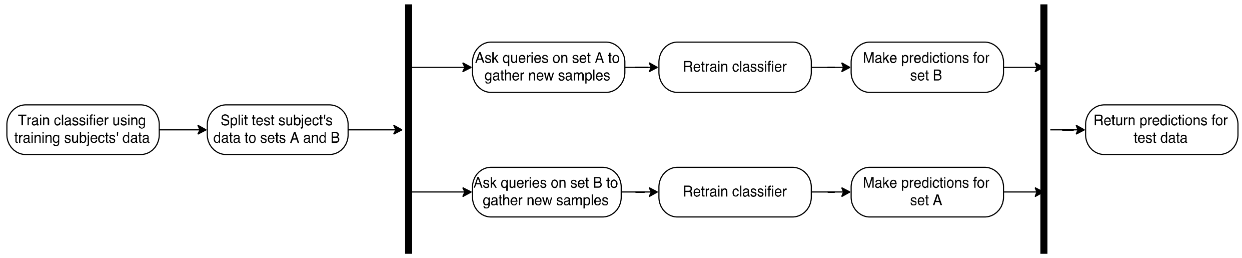

Fig. 3. Workflow of the active learning model evaluation.

C. Comparison of Features

In the first part of the experiment, we trained a supervised model on our dataset to be used as our baseline activity recognition model and compared models that use different combination of features from different chosen sensors.

The following hardware and software-based sensors and combinations of them were examined for accuracy using the activity classification model: accelerometer, linear acceleration sensor, and rotation vector sensor. Features extracted from these sensors were used as an input into 5 different classification algorithms described in [25] as previously utilized in activity recognition problems: Random Forest, Extra Trees, Naïve Bayes, Logistic Regression, and Support Vector Machine. The performance of each model was then tested using leave-one-subject-out (LOSO) cross-validation. Specifically, for each participant, we trained the classifier on the data collected from the other participants and then tested it on the subject's own data. This method was used over 10-fold cross-validation, as cross-validation can obtain deceptively high accuracies in activity recognition classifiers [40].

D. Active Learning Model

In this study, we employed the two most reliable Active Learning algorithms, Uncertainty Sampling and Query by Committee, based on the review conducted in [28], and customized them based on the application needs. These two algorithms allowed us to perform online active learning and query subjects as sensor-data was being received. We compared the performance of these two methods for activity recognition to determine which method is most appropriate for this application. As illustrated in Fig. 3, both Active Learning algorithms queried for labels on two different datasets. It then retrained the classifier for each dataset, and made predictions for the opposite dataset that was not queried. Using this approach, annotating one sample did not result in a falsely high accuracy of the model by having prior knowledge of its future classifications.

1) Uncertainty Sampling

As described earlier, the Uncertainty Sampling strategy selected the most informative samples to be queried for label based on an uncertainty measure. Here, we used a variant of multi-class uncertainty measure, margin sampling which is described in [41]. In this approach, the model was considered uncertain about a sample when the difference of the most probable and second most probable label for a sample was not large enough. Moreover, we defined the probability of a label as the number of estimators in the Extra Trees classifier [42] that predicted the same label divided by the total number of estimators. As a result, the certainty of the prediction was computed as:

$$p(y|x) = \frac{\#estimators\ predicted\ y\ for\ input\ x}{\#\ extimators\ in\ Extra\ Tree\ Classifier} \quad (3)$$

If the algorithm was not certain between two or more labels and had certainty difference less than a threshold, then the subject was queried for the real activity label through the smartwatch interface. This threshold was determined for each activity in a flexible manner. To assign thresholds to each activity type, which prevented skewing of query requests toward specific activities and issuing excessive queries for each activity, we set thresholds so that the percentage of misclassified instances for each activity was the same. Previous samples and labels received by the system gave us an insight into which percentage of samples was expected to be classified correctly, for different given thresholds. With this statistical data, we first determined the percentage of misclassification that was tolerable and the percentage of queries that should be issued. Then, based on these results, a different uncertainty threshold was calculated for each activity. With this approach, when the certainty of an activity was lower, the number of queries issued for that activity increased.

2) Query by Committee

In this study, to implement the Query by Committee strategy in the smartwatch activity recognition system, we needed to define the committee and voting system in a way that allowed the model to support stream-based query generation and decision making based on the current incoming sample. We used 3 different classifier models to determine when to query the individual: Extra Trees classifier, Support Vector Machine with a linear kernel, and a Naïve Bayes classifier. We issued a query when any pair of these classifiers predicted different labels for the incoming data point. After issuing the queries, we then re-trained the classifiers and returned a final decision of the algorithm.

V. Results

A. Baseline Results

Fig. 4 depicts the accuracy of the tested classifiers for each set of sensor combinations. As seen in the figure, the addition of the linear acceleration sensor slightly increased the model's accuracy compared to the accelerometer sensor. More importantly, the combination of the rotation vector sensor with the linear acceleration sensor gave the highest accuracy across all classifiers tested. However, the rotation vector sensor itself did not provide accurate results. Finally, as Fig. 4 illustrates, combining the data from the rotation vector sensor and the accelerometer sensor did not provide significant improvements in accuracy. This may be due to the fact that the orientation of the device is already considered in the accelerometer data analysis.

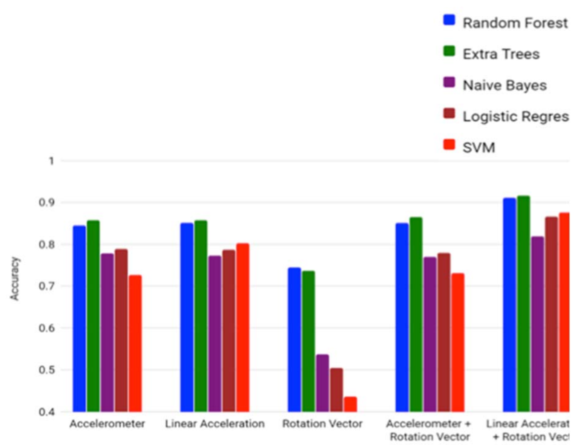

Fig. 4. Accuracy of the classifier tested for each sensor combination.

Fig. 4 also illustrates that the Extra Trees classifier outperformed the other classifiers tested in this application. The hyper-parameters determined for the Extra Trees classifier was 1000 and 5 for the number of estimators and split threshold, respectively.

By selecting the Extra Trees classifier as the main underlying activity model, we were able to compare how well each combination of sensors distinguished activities. Fig. 5 depicts the performance of the Extra Trees classifier given the different sets of sensors using the F1 score.

It can be inferred from the results presented previously that the linear acceleration sensor can accurately classify walking and running activities, however, this sensor performs poorly when detecting standing and sitting activities. Conversely, the combination of linear acceleration and rotation vector sensors resulted in a better or equal performance in classification across all activities compared to any other combinations of sensors. This may be due to the fact that by having access to only the linear acceleration sensor, the algorithm had no sense of the orientation of the coordinate system of the device, which was necessary for classification of certain activity types. Overall, the software-based sensors used in the model, linear acceleration and the rotation vector sensor, improved the overall accuracy of the supervised activity recognition model from approximately 85% to 93% when the Extra Trees classifier was used. We used this model as our baseline classifier for evaluating the active learning model as it produced the highest performances compared to the other classifiers tested. The average accuracy achieved for each subject using this model is provided in Table 2. These results demonstrate that the current smartwatch activity recognition system is a good option for this type of application.

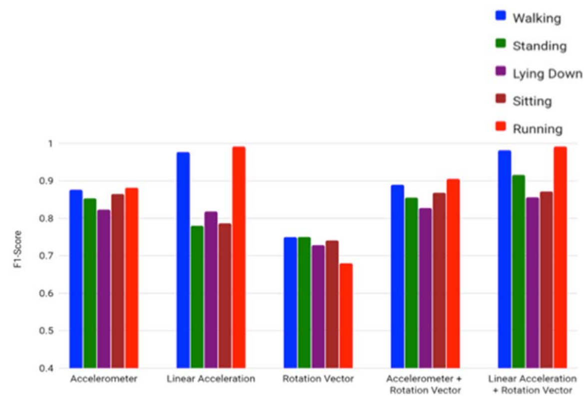

Fig. 5. F1 score of activities obtained by the Extra Trees classifier for each sensor combination.

TABLE II. F1 Score Across Subjects For the Baseline model

Subject ID	Baseline Accuracy	Subject ID	Baseline Accuracy
1	0.985	**7**	0.950
2	0.893	**8**	0.943
3	0.946	**9**	0.957
4	0.935	**10**	0.839
5	0.843	**11**	0.858
6	0.954	**12**	0.991

B. Active Learning Resutls

1) Uncertainty Sampling

Fig. 6 shows the cumulative distribution of certainty among correctly and incorrectly classified instances for each activity type. In other words, for each level of certainty, the graphs show the percentage of correctly and incorrectly classified instances with less or equal certainty. As seen in this figure, the classifier had less certainty for misclassified instances which is expected. Therefore, by querying for uncertain points, we gain information and decrease the uncertainty for that particular activity.

Another interesting result that can be seen in Fig. 6 is that in each threshold, a different percentage of misclassification occurs for different activities. This can be used to balance number of queries for each activity. For instance, when the certainty threshold is set to 60%, a data point that is incorrectly predicted as walking is a good candidate to be annotated by the user for the true label. Specifically, this threshold is less than

10% of the data points that were correctly classified as walking while the threshold for lying down is 40%. Therefore, by determining the percentage of misclassified instances and consequently calculating the threshold for each activity type, the model is able to decide whether to query for a label or not. For instance, to query for 80% of misclassified instances, we obtain a certainty threshold of 61% for walking and 42% for standing.

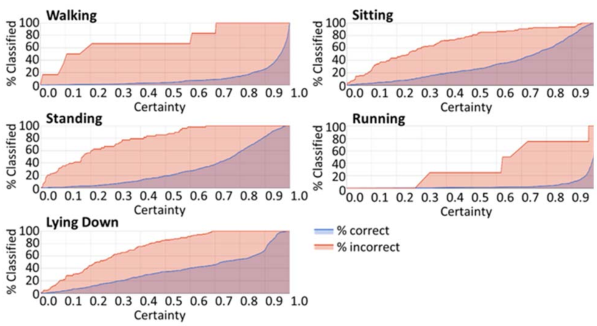

Fig. 6. A comparison of the F1 score for each activity achieved by different Active Learning approaches.

Fig. 7 shows the results of our active learning model using the Uncertainty Sampling algorithm for each subject. It can be determined from this figure, particularly from the increase in accuracy using different thresholds, that the two subjects with the poorest classified activities in the baseline model had a considerable increase in accuracy after implementing the Active Learning model. Specifically, there was more than a 7% increase in accuracy of activity classification for Subject 9. This is due to the personalization of the model that is inherent in the Active Learning algorithm. Moreover, after assessing the accuracy of classifying each activity, depicted in Fig. 8, our model improved activities that were particularly harder to distinguish from other activities such as standing and sitting. This is true given that the number of training samples was 46% of the total number of samples available in the dataset.

Another important result found from this analysis is that the number of queries issued for each subject was correlated with the amount of increase in accuracy that the active learning model achieved. For example, when the active uncertainty threshold was set to 60%, our algorithm issued only 21 queries for Subject 1, which had the most accurate predictions, while it issued 91 queries for Subject 9, which had the least accurate predictions in the baseline model.

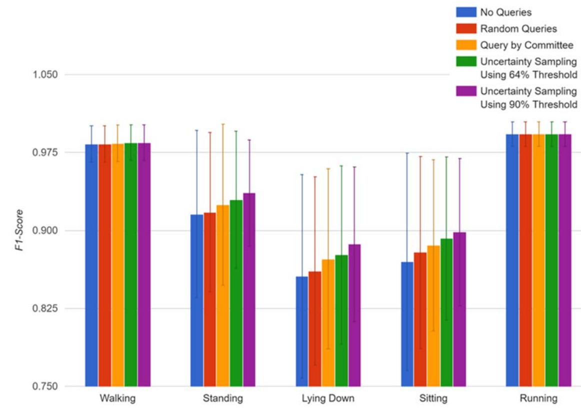

Fig. 8. Model certainty based on the number of visited samples for the different activities.

C. *Query by Committee Results*

Fig. 8 also compares the mean and standard deviation of the F1 score achieved for each query strategy. To objectively compare the performance of different query strategies, we issued the same number of queries. As a result, the Query by Committee strategy issued 566 total queries, and the Uncertainty Sampling strategy with a threshold set to 64% issued the same number of queries. Moreover, a random query generator was used as baseline to compare the performance across these two methods. This random query generator was also designed such that it generated the same total number of queries.

It can be seen that both strategies outperformed the random query strategy. More importantly, given the same number of queries, the Uncertainty Sampling strategy was slightly better than the Query by Committee strategy. The Uncertainty Sampling strategy is also more flexible for activity recognition problems as it can modify the query issuance rate by modifying the algorithm uncertainty threshold. As it can be seen in Fig. 9,

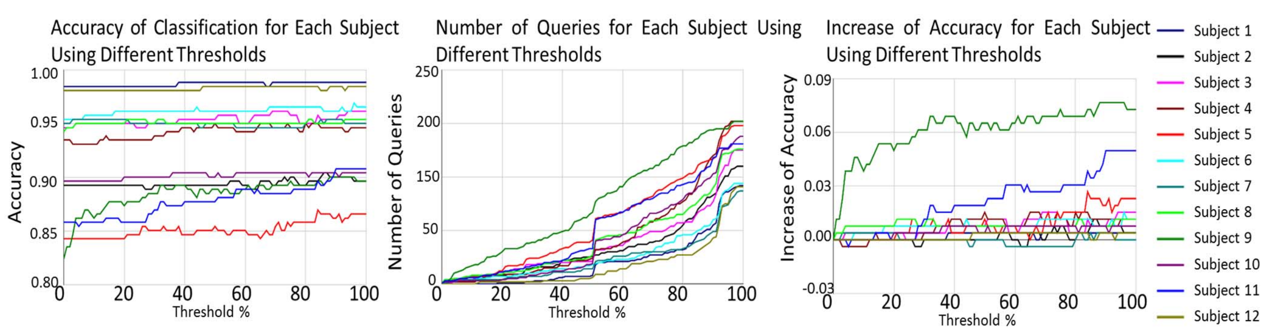

Fig. 7. Simulation results after performing Uncertainty Sampling. Accuracy, number of queries, and increase in accuracy for each subject is shown based on the uncertainty threshold.

using a 90% threshold in the Uncertainty Sampling strategy completely outperformed the other strategies, even after issuing

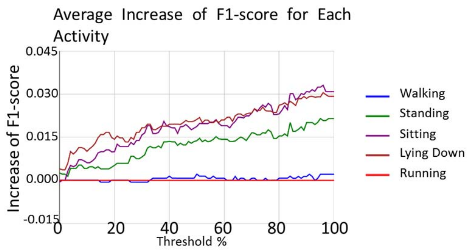

Fig. 9. Average increase of F1-score for each activity using the active learning model.

queries for all samples. Table 3 provides the final accuracy that is achieved by our Active Learning model for each subject.

By comparing results of Table 3 with Table 2, we can see that our approach improved the activity recognition accuracy across all subjects as a result of the personalization capabilities in our approach. Furthermore, our approach used far less number of samples for training, further making it ideal for real-world activity recognition problems.

TABLE III. F1 Score Across Subjects For the Active Learning Model.

Subject ID	Active Learner Accuracy	Subject ID	Active Learner Accuracy
1	0.989	**7**	0.954
2	0.900	**8**	0.950
3	0.954	**9**	0.900
4	0.946	**10**	0.904
5	0.962	**11**	0.908
6	0.962	**12**	0.985

VI. Discussion

In this study, we provided a novel approach for activity recognition problems through the use of smartwatch-based active learning methods. By including the most recent advancement of technology in our method, we were able to significantly increase the accuracy of activity recognition classifiers and provide personalized activity models for each individual. Finally, we showed that recent advancements in smartwatches and their increase in popularity among the general public enables future applications to use them as an accurate and usable activity recognition tool.

We demonstrated that using two software-based sensors, namely the rotation vector sensor and linear acceleration sensor, the performance of activity classifiers can be improved. We then approached the problem of ground truth creation by providing methods for labeling samples that is feasible and reduces user burden in real world activity recognition problems. Labeling samples is cumbersome, or in some cases, problematic, particularly during activity recognition applications that target elderly, disabled or pediatric populations. We showed that through the use of active learning machine learning models, we can utilize not only a smaller number of samples for dataset creation, but also create personalized models for each individual. Furthermore, through the use of active learning algorithms, we were able to achieve an average accuracy of 92% using 46% less samples than supervised machine learning methods.

Future work will focus on developing new stream-based active learning algorithms that can update the model quickly and on the fly. It is also desirable to have light algorithms that enable us to update the model on the smartwatch while reducing the amount of power consumption. Furthermore, noise and outlier detection algorithms for streams of data will be improved and applied to our smartwatch-based active learning algorithm. The query strategies described here will also be further developed so that they consider the number of queries made to reduce user burden and improve compliance with use of the system.

VII. Conclusion

In this study, we demonstrated that an active learning approach for activity recognition using sensor data collected from smartwatches is a viable option for future applications. Given the desirable personalized models that result from this method, this approach can be applicable to all types of populations that require activity monitoring. The high accuracy we achieved in this study given the small number of required training samples demonstrated that this approach is significantly better than previously studied supervised learning methods. Finally, we demonstrated that the use of software-based smartwatch sensors can improve the accuracy of activity recognition models and provide a desirable platform for future applications.

References

[1] C.C. Yang, and Y.L. Hsu. "A review of accelerometry-based wearable motion detectors for physical activity monitoring," Sensors, vol. 10(8), pp. 7772-7788, 2010.

[2] K. Taraldsen, S.F. Chastin, I.I. Riphagen, B. Vereijken, and J.L. Helbostad. "Physical activity monitoring by use of accelerometer-based body-worn sensors in older adults: a systematic literature review of current knowledge and applications," Maturitas, vol. 71(1), pp. 13-19, 2012.

[3] J. Ko, C. Lu, M. B. Srivastava, J. A. Stankovic, A. Terzis and M. Welsh, "Wireless Sensor Networks for Healthcare," in *Proceedings of the IEEE*, vol. 98, no. 11, pp. 1947-1960, Nov. 2010.

[4] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu and P. Havinga, "Activity Recognition Using Inertial Sensing for Healthcare, Wellbeing and Sports Applications: A Survey,"*Architecture of Computing Systems (ARCS), 2010 23rd International Conference on*, Hannover, Germany, pp. 1-10, 2010.

[5] N. Amini, M. Sarrafzadeh, A. Vahdatpour, and W. Xu. "Accelerometer-based on-body sensor localization for health and medical monitoring applications". *Pervasive and mobile computing*, vol. 7, no. 6, pp.746-760, 2011.

[6] A. Vahdatpour, N. Amini, and M. Sarrafzadeh. "On-body device localization for health and medical monitoring applications," in *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*, pp. 37-44, March 2011.

[7] M.K. Suh, C.A. Chen, J. Woodbridge, M.K. Tu, J.I. Kim, A. Nahapetian, L.S. Evangelista, and M. Sarrafzadeh. "A remote patient monitoring system for congestive heart failure," *Journal of medical systems*, vol. 35, no. 5, pp.1165-1179, 2011.

[8] O.D. Lara and M.A. Labrador, "A Survey on Human Activity Recognition using Wearable Sensors," in *IEEE Communications Surveys & Tutorials*, vol. 15, no. 3, pp. 1192-1209, Third Quarter 2013.

[9] J.R. Kwapisz, G.M. Weiss, and S.A. Moore. "Activity recognition using cell phone accelerometers." *ACM SigKDD Explorations Newsletter* vol. 12.2, pp. 74-82, 2011.

[10] A.M. Khan, Y.K. Lee, S.Y. Lee and T.S. Kim, T. S. "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis." *2010 5th International Conference on Future Information Technology*. IEEE, 2010.

[11] D. Anguita, 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." *International Workshop on Ambient Assisted Living*. Springer Berlin Heidelberg, 2012.

[12] D. Anguita, A. Ghio, L. Oneto, X. Parra and J.L. Reyes-Ortiz. "A Public Domain Dataset for Human Activity Recognition using Smartphones." *ESANN*. 2013.

[13] Y.S. Lee and S.B. Cho. "Activity recognition using hierarchical hidden markov models on a smartphone with 3D accelerometer."*International Conference on Hybrid Artificial Intelligence Systems*. Springer Berlin Heidelberg, 2011.

[14] M. Shoaib, S. Bosch, H. Scholten, P.J. Havinga, and O.D. Incel. "Towards detection of bad habits by fusing smartphone and smartwatch sensors." *Pervasive Computing and Communication Workshops (PerCom Workshops), 2015 IEEE International Conference on*. IEEE, 2015.

[15] R. Rawassizadeh, B.A. Price, and M. Petre. "Wearables: has the age of smartwatches finally arrived?." *Communications of the ACM*, vol. 58.1, pp. 45-47, 2015.

[16] Y. Zhu, E. Hunt, K. Magsamen-Conrad, and J.M. Bernhardt. "Facilitators and barriers to smartwatch use among individuals with chronic diseases: A qualitative study", Elizabeth M. Glowacki University of Texas at Austin.

[17] G. Phillips, L. Felix, L. Galli, V. Patel, and P. Edwards. "The effectiveness of m-health technologies for improving health and health services: a systematic review protocol", BMC Research Notes, vol. 3, no. 1, pp. 250, 2010.

[18] G.F. Dunton, Y. Liao, S.S. Intille, D. Spruijt-Metz, and M. Pentz. "Investigating children's physical activity and sedentary behavior using ecological momentary assessment with mobile phones", Obesity, vol. 19, no. 6, pp. 1205-12212, 2011.

[19] Z.S. Abdallah, M.M. Gaber, B. Srinivasan, and S. Krishnaswamy. "Adaptive mobile activity recognition system with evolving data streams." *Neurocomputing,* vol. 150, pp. 304-317, 2015.

[20] L. Bao and S.S. Intille. "Activity recognition from user-annotated acceleration data." *International Conference on Pervasive Computing*. Springer Berlin Heidelberg, 2004.

[21] S. Pirttikangas, K. Fujinami, and T. Nakajima. "Feature selection and activity recognition from wearable sensors." *International Symposium on Ubiquitious Computing Systems*. Springer Berlin Heidelberg, 2006.

[22] J. Parkka, M. Ermes, P. Korpipaa, J. Mantyjarvi, J. Peltola, and I. Korhonen. "Activity classification using realistic data from wearable sensors." *IEEE Transactions on information technology in biomedicine,* vol. 10.1, pp. 119-128, 2006.

[23] L. Atallah, B. Lo, R. King, and G.Z. Yang. "Sensor positioning for activity recognition using wearable accelerometers." *IEEE transactions on biomedical circuits and systems*, vol. 5.4, pp. 320-329, 2011.

[24] M. Keally, G. Zhou, G. Xing, J. Wu, and A. Pyles. "Pbn: towards practical activity recognition using smartphone-based body sensor networks." *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2011.

[25] I. Cleland, B. Kikhia, C. Nugent, A. Boytsov, J. Hallberg, K. Synnes, S. McClean, and D. Finlay. "Optimal placement of accelerometers for the detection of everyday activities." *Sensors* vol. 13.7, pp. 9183-9200, 2013.

[26] D. Guan, W. Yuan, Y.K Lee, A. Gavrilov, and S. Lee. "Activity recognition based on semi-supervised learning." 13th IEEE International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA 2007). IEEE, 2007.

[27] D. Angluin. "Queries and concept learning." *Machine learning,* vol. 2.4, pp. 319-342, 1988.

[28] B. Settles. "Active learning literature survey." *University of Wisconsin, Madison*, vol. 52.55-66, pp. 11, 2010.

[29] H.S. Seung, M. Opper, and H. Sompolinsky. "Query by committee." *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, 1992.

[30] D.D. Lewis, and W.A. Gale. "A sequential algorithm for training text classifiers." *Proceedings of the 17th annual international ACM SIGIR conference on Research and development in information retrieval*. Springer-Verlag New York, Inc., 1994.

[31] T. Miu, P. Missier, and T. Plötz. "Bootstrapping Personalised Human Activity Recognition Models Using Online Active Learning." *Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), 2015 IEEE International Conference on*. IEEE, 2015.

[32] W. Albert and T. Tullis. "Measuring the user experience: collecting, analyzing, and presenting usability metrics". Newnes, 2013.

[33] X. Su, H. Tong, and P. Ji. "Activity recognition with smartphone sensors." *Tsinghua Science and Technology*, vol. 19.3, pp. 235-249, 2014.

[34] A.M. Otebolaku and M.T. Andrade. "Recognizing high-level contexts from smartphone built-in sensors for mobile media content recommendation." *2013 IEEE 14th International Conference on Mobile Data Management*, vol. 2. IEEE, 2013.

[35] J.R. Dooley and J.M. McCarthy. "Spatial rigid body dynamics using dual quaternion components." *Robotics and Automation, 1991. Proceedings., 1991 IEEE International Conference on*. IEEE, 1991.

[36] D. Lewis and W. Gale. A sequential algorithm for training text classifiers. *Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 3–12, 1994.

[37] H.S. Seung, M. Opper, and H. Sompolinsky. Query by committee. *Proceedings of the ACM Workshop on Computational Learning Theory*, pp. 287–294, 1992.

[38] J. Wan, M.J. O'Grady, and G.M.P. O'Hare. "Dynamic sensor event segmentation for real-time activity recognition in a smart home context." *Personal and Ubiquitous Computing,* vol. 19.2, pp. 287-301, 2015.

[39] J. Yang. "Toward physical activity diary: motion recognition using simple acceleration features with mobile phones." *Proceedings of the 1st international workshop on Interactive multimedia for consumer electronics*. ACM, 2009.

[40] A. Stisen, H. Blunck, S. Bhattacharya, T.S. Prentow, M.B. Kjærgaard, A. Dey, T. Sonne, and M.M. Jensen. "Smart devices are different: Assessing and mitigatingmobile sensing heterogeneities for activity recognition." *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, ACM, 2015.

[41] T. Scheffer, C. Decomain, and S. Wrobel. "Active hidden markov models for information extraction." *International Symposium on Intelligent Data Analysis*, Springer Berlin Heidelberg, 2001.

[42] P. Geurts, D. Ernst, and L. Wehenkel. "Extremely randomized trees." *Machine learning,* vol. 63.1, pp. 3-42, 2006.