

Feature Engineering for Human Activity Recognition

Basma A. Atalaa^{1*}, Ibrahim Ziedan², Ahmed Alenany³, Ahmed Helmi⁴

Department of Computer and Systems Engineering

Faculty of Engineering

Zagazig University, Zagazig, 44519

Egypt

Abstract—Human activity recognition (HAR) techniques can significantly contribute to the enhancement of health and life care systems for elderly people. These techniques, which generally operate on data collected from wearable sensors or those embedded in most smart phones, have therefore attracted increasing interest recently. In this paper, a random forest-based classifier for human activity recognition is proposed. The classifier is trained using a set of time-domain features extracted from raw sensor data after being segmented into windows of 5 seconds duration. A detailed study of model parameter selection is presented using the statistical *t*-test. Several simulation experiments are conducted on the WHARF accelerometer benchmark dataset, to compare the performance of the proposed classifier to support vector machines (SVM) and Artificial Neural Network (ANN). The proposed model shows high recognition rates for different activities in the WHARF dataset compared to other classifiers using the same set of features. Furthermore, it achieves an overall average precision of 86.1% outperforming the recognition rate of 79.1% reported in the literature using Convolution Neural Networks (CNN) for the WHARF dataset. From a practical point of view, the proposed model is simple and efficient. Therefore, it is expected to be suitable for implementation in hand-held devices such as smart phones with their limited memory and computational resources.

Keywords—Human activity recognition; random forest; feature engineering; sensor signal processing

I. INTRODUCTION

In daily life, a person performs diverse set of activities such as standing up, sitting down, walking, climbing stairs, etc. Automatic recognition of human activities has interesting applications in healthcare [1], keeping track of elderly people [2], and home automation [3]. Also, it has many clinical applications for stroke patients [4], Parkinson's disease patients [5], heart rate estimation [6] and in a smart health care environment [7].

The last two decades witnessed increasing interest in Human Activity Recognition (HAR) techniques due to the availability of low cost sensors specially those built-in sensors available in affordable smartphones [8-10]. Commonly used sensor types in HAR applications are accelerometers [11-14], heart rate belt sensor [15], gyroscope [16, 17], magnetometer [17], or three-inertial sensor units mounted on chest, right thigh and left ankle [12]. Such inertia devices operate at low frequencies and require low sampling rates. There are several issues which make HAR task challenging such as noisy sensor data, insufficient training examples due to few participating subjects, and the need to implement HAR systems on

relatively limited-resources smart devices. Therefore, numerous studies in literature have been conducted to look for suitable representative features for activities, as well as good enough recognition models [9]. Moreover, benchmark datasets available in literature are different in type of activities, number of recorded examples for each activity, experimental settings, i.e. controlled procedure [18] whether indoor or outdoor environments [19], used sensors and sensor position on subject body. According to aforementioned factors, there is a significant variance of available HAR systems accuracy in conjunction with different datasets [20].

HAR recognition techniques can be grouped into two main categories. The first is based on computer vision [21, 22] and the second is based on data collected from one or more sensors. What makes the latter approach appealing is that sensors are affordable and are usually found in reasonably priced smartphones. Another advantage is that computational and storage requirements for processing sensor data is less than those required for image processing techniques.

In this work, the relatively challenging Wearable Human Activity Recognition Folder (WHARF) dataset is extensively investigated. This dataset is collected using a tri-axial accelerometer placed on the right wrist of subjects; hence it emulates a smart watch. It is challenging because of its small sampling rate, 32 Hz, compared to other datasets collected using e.g. 50 Hz sampling frequency. Real-time considerations for HAR systems require dealing with segments of data points with window length between 2 seconds and 10 seconds. Therefore, sensors with small sampling rate will deliver fewer data points complicating the task of HAR system. Moreover, there are 12 different activities in WHARF with few number of examples per activity [13]. The proposed approach here applies data preprocessing in which signals are filtered using a low-pass filter and then scaled so that all features lie within the same range. In the second step, data is segmented into windows of length 5 seconds with 50% overlapping. In the third step, several effective time-domain functions or features are extracted. The proposed classifier employs the Random Forest (RF) algorithm which achieves the best precision and also the best training time compared to other classifiers such as Artificial Neural Networks (ANN) and Support Vector Machine (SVM). The proposed system is expected to be efficient and resource-friendly for smart devices. Besides, sensitivity analysis of proposed system components such as RF parameters, some important features and preprocessing scaling step is conducted. Also, feature importance is discussed using the statistical *t*-test.

*Corresponding Author

The contribution of this work can be highlighted as follows: (1) introducing RF-based effective and efficient HAR system with average precision of 86.1% and average accuracy of 84.8% which improves the state-of-the-art rate of 79.1% for WHARF dataset, (2) testing the proposed system on the challenging WHARF dataset which is considered in only few studies in literature [23] and [24], (3) discussing the practical implementation issues of proposed system which is important in case of further system application on smart devices, and (4) conducting sensitivity analysis of important system components to determine the optimal settings for proposed system.

The rest of this paper is organized as follows. In Section II, relevant related work in the literature is reviewed. The set of features to be employed and the proposed Random Forest-based classifier are presented in Sections III and IV, respectively. In Section V, a set of experiments are conducted to evaluate the performance of the proposed model and compare it to other machine learning techniques. Sensitivity analysis is performed to optimally select the parameters of the proposed model in Section VI. Finally, conclusions and possible future work are drawn in Section VII.

II. RELATED WORK

The HAR procedure from preprocessed raw sensory data can be divided into two steps: (1) extracting relevant key features from collected data signals (so-called feature engineering), and (2) classifying the observed activity based on the extracted features. The reduction of data dimensionality may also be required using e.g. principle component analysis [25]. Due to the diversity of feature types and the classifiers that can be used in these two steps, respectively, the literature of HAR problem is wide and extensive.

Sensors such as tri-axial accelerometer and gyroscope provide time domain acceleration and angular velocity readings in the x , y , and z axes, respectively. In the literature, the various types of features which are extracted from such raw data can be divided into two categories:

1) *Time domain features*: e.g. the coefficients of an autoregressive (AR) model for each of the x , y , and z axes [11, 18, 26-29], signal magnitude area (SMA) [11, 18, 26-28, 30], tilt angle [11, 31], Histogram [17], mean [17, 26, 31], standard deviation [25, 26], Jerk [32, 33], roll angle [11, 24] skewness, kurtosis and total integral of modulus of accelerations (IMA) [12], and.

2) *Frequency domain features*: e.g. power spectral density (PSD) [12, 25], signal entropy and spectral energy [12, 31], largest frequency component, average frequency signal skewness, and frequency signal kurtosis [26].

It should be noted that the use of various types of features is important to improve the classification task. Each class of activities has its own set of discriminative features which is in general different from other classes. For example, the standard deviation feature can be used to distinguish between static and dynamic activities, and the Fast Fourier transform (FFT) coefficients can be used to distinguish between walking and running [11].

On the other hand, classifiers used in HAR studies can be classified into supervised or unsupervised. Supervised classifiers [20] include multilayer neural networks [17, 18, 30, 31, 34], support vector machine (SVM) [11, 12], decision trees [30, 31], random forest [12], k-Nearest Neighbors (kNN) [12, 16] and Bayes classifier [16, 25]. Unsupervised technique, on the other hand, include Gaussian mixture model (GMM) [13], linear-discriminant analysis [27, 28], minimal learning machine (MLM) [16], k -means clustering, convolutional neural networks (CNN) [35-37] and hidden Markov model (HMM) [12].

III. TIME-DOMAIN AND STATISTICAL FEATURES

In this section, the set of features extracted from pre-processed raw acceleration signals is listed. It is assumed that there is a three-dimensional dataset of size N data points collected from an accelerometer or a gyroscope, $a_x(i)$, $a_y(i)$, $a_z(i)$, $i=1, 2, \dots, N$, for the x , y , and z dimensions. The data is first filtered using low pass filter to reduce noise and extract the body acceleration $b_x(i)$, $b_y(i)$, $b_z(i)$ and gravity acceleration $g_x(i)$, $g_y(i)$, $g_z(i)$ components [24].

The set of features to be employed in classification are derived from both body and gravity acceleration signals as listed in Table I. The body acceleration signal features include the mean (M) and standard deviation (STD) of filtered signals, autoregressive model coefficients, signal magnitude area, tilt angle, mean, standard deviation, entropy of jerk of signals, mean, standard deviation, power and entropy of jerk of roll angle. For gravity acceleration component, the signal power along each axis and the mean of angle of x -axis component are used.

IV. THE PROPOSED MODEL

The proposed classifier consists of three stages as shown in Fig. 2. In the first stage, the data is applied to a low pass filter to filter out noise and separate body acceleration from gravity acceleration. The data is then segmented into windows of 5 seconds duration consisting of 160 data points. In the second stage, the set of features listed in Table I are extracted. Finally, the classification task is performed in the third stage using random forest classifier [12].

Random Forest can be described as an ensemble or set of decision trees as shown in Fig. 2 where each tree produces a prediction of the class to which the given example belongs. The overall decision is then made using a voting process on the most predicted class among all trees in the forest. Random forest classifier has several so-called hyper-parameters which affect the classification. These include the number of trees in the forest and the maximum depth of the trees. The default value for number of trees is 100 whereas the default value for the maximum depth is 0. This means that each tree will expand until every leaf is pure, i.e. all data on the leaf comes from the same class. Random Forest classifier first selects random feature vectors from the dataset, builds a decision tree for each sample and performs a vote to determine the most voted prediction. In the current work, the basic RF classifier is employed in HAR recognition. To find the optimal RF parameters, a sensitivity analysis is conducted in Section

TABLE I. LIST OF FEATURES AND THEIR FORMULAS

| Term | Meaning | Formula | Scaling factor |
|--|---|---|--|
| Autoregressive (AR) model coefficients | Autoregressive model is used to predict time series data from past data records in x , y and z -directions | $b_x(n) = -\sum_{k=1}^p a(k) b_x(n-k) + e(n)$ | $1/\sqrt{\ b_x\ }$ |
| Signal magnitude area | A scalar feature used to distinguish static from dynamic activities such as standing and walking [11] | $SMA = \frac{1}{N} \sum_{i=1}^N (b_x(i) + b_y(i) + b_z(i))$ | $1/\ (b_x)^2\ $ |
| Tilt Angle | Angle between z -axis and gravitational vector g . It is used to distinguish postures such as standing and lying [11] | $\phi = \frac{1}{N} \sum_{i=1}^N \arcsin\left(\frac{b_z(i)}{\ b_z\ }\right)$ | $1 \times \sqrt{\ b_x\ }$ |
| Jerk | The rate of change of body acceleration. | $J = \frac{\partial b_x(i)}{\partial t}$ | - |
| Roll angle | Describes the rotation of accelerometer attached to the participant's hand about x -axis as shown in Fig. 1 [24] | $\phi(t) = \tan^{-1}(-b_z(t), -b_y(t))$ | $1/\sqrt{\ b_x\ }$ |
| Angle of x -axis gravity signal | This angle is used to estimate sensor attitude | $\theta = \text{real}\left(\cos^{-1}\left(\max\left(\min\left(\frac{u^T v}{\ u\ \ v\ }, 1\right), -1\right)\right)\right)$ | $1/\ b_x\ $ |
| Power | Signal power | $P_x = \left(\sqrt{\frac{1}{N} \sum_{i=1}^N (g_x(i)^2)}\right)^2$ | - |
| Entropy of signal (S) | Statistical measure of signal randomness | $E(S) = -\sum_{i=1}^n (P(s_i) \log_2 P(s_i))$ | - |
| Mean | Describes the central tendency or the dc level of the signal | $\mu_x = \frac{1}{N} \sum_{i=1}^N b_x(i)$ | $1/\sqrt{\ b_x\ }$ for b_x, b_y, b_z and jerk |
| Standard deviation | Describes the amount of variation around the mean | $\sigma_x = \sqrt{\frac{1}{N} \sum_{i=1}^N (b_x(i) - \mu_x)^2}$ | $1/\sqrt{\ b_x\ }$ for b_x, b_y, b_z |

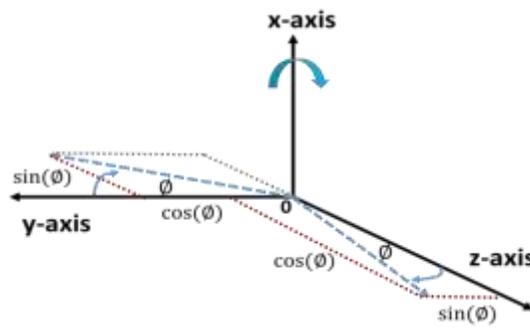
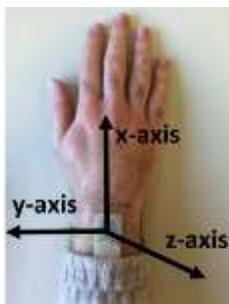


Fig. 1. Accelerometer Orientation during WHARF Dataset Collection [23] and (b) Roll Angle (ϕ) after Rotation Around x -axis.

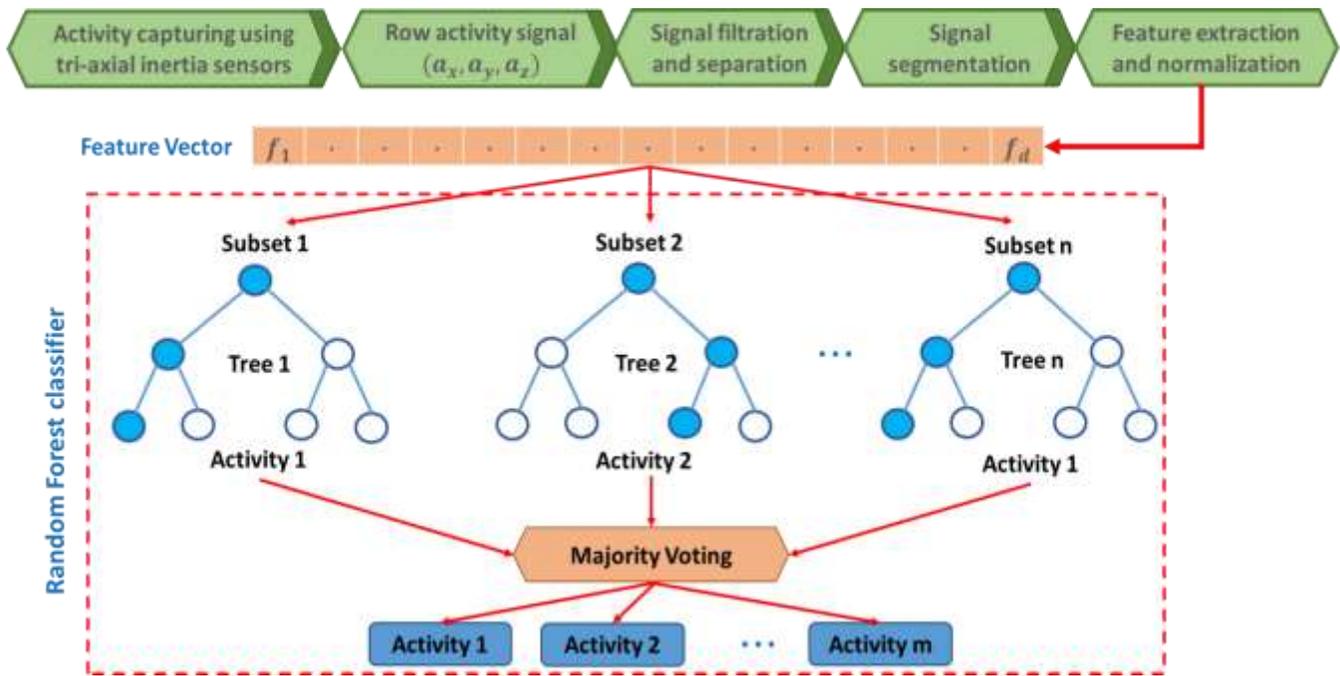


Fig. 2. Block Diagram of the Proposed Human Activity Recognition System.

V. EXPERIMENTAL RESULTS

A. Dataset

In this section, the benchmark Wearable Human Activity Recognition Folder (WHARF) dataset by Bruno et al. [13], is used to examine the performance of the proposed HAR technique. The dataset was collected by an ad-hoc tri-axial accelerometer sensor attached to the right wrist of the participant. The participants are 17 volunteers; 11 males, with age ranging from 19 to 81 years; and 6 females, with ages between 56 and 85 years [11]. The digital resolution of the sensor is 6 bits and the sampling rate is 32 Hz. The dataset contains the following 12 activities: Brush_teeth (BT), Climb_stairs (CS), Comb_hair (CH), Descend_stairs (DS), Drink_glass (DG), Getup_bed (GB), Liedown_bed (LB), Pour_water (PW), Sitdown_chair (SD), Standup_chair (SU), Use_telephone (UT) and Walk (WK). The examples of each activity class are contained in a separate folder and raw signals for each single activity are saved in one text file.

B. Classification Rates

According to recent studies in the literature [23, 26, 35], classification results of different classifiers and settings have been reported in terms of the Precision (or positive predictive rate) and the Recall (or sensitivity) as the most crucial metrics in HAR applications. Let TP, FP and FN denote true positive, false positive and false negative, respectively, then the precision (P) can be calculated as $P = \frac{TP}{TP+FP}$, whereas the recall (R) is expressed as $R = \frac{TP}{TP+FN}$.

All experiments were conducted using machine learning package Sklearn in Python. Each activity signal is segmented into windows of 5 seconds duration [24] in order to fulfil real-world demands of HAR systems [26]. In Table II, a comparison is made between the proposed model using random forest against SVM and ANN. The results show that SVM and ANN have better precision than random forest in some activities. For example, SVM achieves 92.1% for Walking while ANN achieves 97% for Descend_stairs activity. However, the proposed model outperforms both SVM and ANN in terms of the average precision achieving 86.1% over all activities.

TABLE II. COMPARISON OF THREE CLASSIFIERS USING THE SAME FEATURE SET IN TERMS OF PRECISION METRIC (%). THE ACTIVITIES ARE BRUSH_TEETH (BT), CLIMB_STAIRS (CS), COMB_HAIR (CH), DESCEND_STAIRS (DS), DRINK_GLASS (DG), GETUP_BED (GB), LIEDOWN_BED (LB), POUR_WATER (PW), SITDOWN_CHAIR (SD), STANDUP_CHAIR (SU), USE_TELEPHONE(UT) AND WALK (WK)

| | BT | CS | CH | DS | DG | GB | LB | PW | SD | SU | UT | WK | Av. Pre. |
|-----|-------------|-----------|-------------|-----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| SVM | 83.1 | 73.8 | 86.3 | 87.8 | 85.3 | 66.4 | 46.2 | 83.6 | 75.6 | 65.4 | 97.3 | <u>92.1</u> | 78.6 |
| ANN | 92 | 74.3 | <u>96.9</u> | <u>97</u> | 88.2 | 63.8 | 68.4 | 79.2 | 79.2 | 64.2 | 82.6 | 82.4 | 80.7 |
| RF | <u>94.6</u> | <u>85</u> | 91 | 94.1 | <u>90.7</u> | <u>75.2</u> | <u>72.2</u> | <u>81.6</u> | <u>88.8</u> | <u>85.1</u> | <u>92.7</u> | 82.4 | <u>86.1</u> |

VI. SENSITIVITY ANALYSIS AND DISCUSSION

A. RF Hyper-Parameters

The hyper-parameters of a random forest, number of trees and maximum tree depth, has a significant effect on the performance of the proposed classifier. To determine the optimal values for these parameters, the classifier is extensively tested using different sets of parameters to obtain the best possible precision. The results of this experiment is shown in Fig. 3 where it can be noticed that as the number of estimators increases, the size of the model on the disk significantly increases, however, without significant increase in the precision. Therefore, the best precision obtained is 86.1% with 100 trees and maximum depth 20. The model size on the disk, in this case, is 16 MB which is reasonable.

B. Effect of Feature Scaling and Normalization

In the proposed model, the use of normalized (scaled) features leads to an average precision of 86.1% as shown in the first row in Table III. Using un-normalized features, however, reduces the average precision to 83.5% as shown in the second row in Table III. This emphasizes the importance of feature normalization.

C. Feature Reduction based on the T-Test

It is important to check the validity and strength of features independently of the classifier to be used afterwards. This simplifies the analysis of the model and reduces the overhead of re-running the whole model several times to check the effect of every feature on the performance.

To evaluate the power of a given feature f_i in discriminating between two classes, the following t -test formula can be used

$$t(f_i) = \frac{|\mu_{i1} - \mu_{i2}|}{\sqrt{\frac{\sigma_{i1}^2}{s_1} + \frac{\sigma_{i2}^2}{s_2}}}$$

where f_i denotes the i -th feature, μ_{i1} and μ_{i2} are the sample means, σ_{i1} and σ_{i2} are the sample standard deviations, and s_1 and s_2 are the size of the two classes, respectively.

Fig. 4 shows the percentage of effective features in each model. For each feature column, a two-sample t -test was carried out between one independent activity class and the other classes in each dataset, is a so-called *one-versus-all* binary classification. It is possible to conclude that if average t_{val} is less than or equal to a critical threshold of 10, the feature is not discriminative enough and could be safely eliminated. It is found that the average t -value for 13 features is less than or equal to 10 and, hence, the size of feature vector reduces to only 24 features. The previous experiments are repeated using the reduced feature vector and the results are shown in Table III. It can be seen from Table III that using the reduced set of features, the average precision decreases from 84%, using full set of features, to only 82% which may not be acceptable.

Table III also shows that the precision is high for some classes and low for others. As can be noticed in Table III, for the activities which made by hand, such as Brush_Teeth, Comb_Hair, Drink_Glass and Use_Telephone, good precision is obtained. Recalling that the sensor is attached to the wrist of the right hand, it can be realized that the position of the sensor helps in capturing these activities in a better way.

D. Size on the Disk and Training Time

In this part, we compare the proposed model using random forest to ANN and SVM in terms of model size in memory and training, and inference times. The results are shown in Table IV where it can be seen that the proposed model is superior to SVM in terms of training time. The proposed model is even better compared to ANN, although the difference is not significant if the number of iterations used in training ANN is reduced. On the other hand, the size of the proposed random forest classifier, 16 MB, is large compared to the other two classifiers and its inference is slower. The large size of the RF classifier is due to the large number of trees, 100, employed. Although the proposed model has large size and slower inference time, which is only 0.01 seconds, they are still reasonable. This combined with fast training time, makes the proposed classifier suitable for use in smartphones and hand held devices.

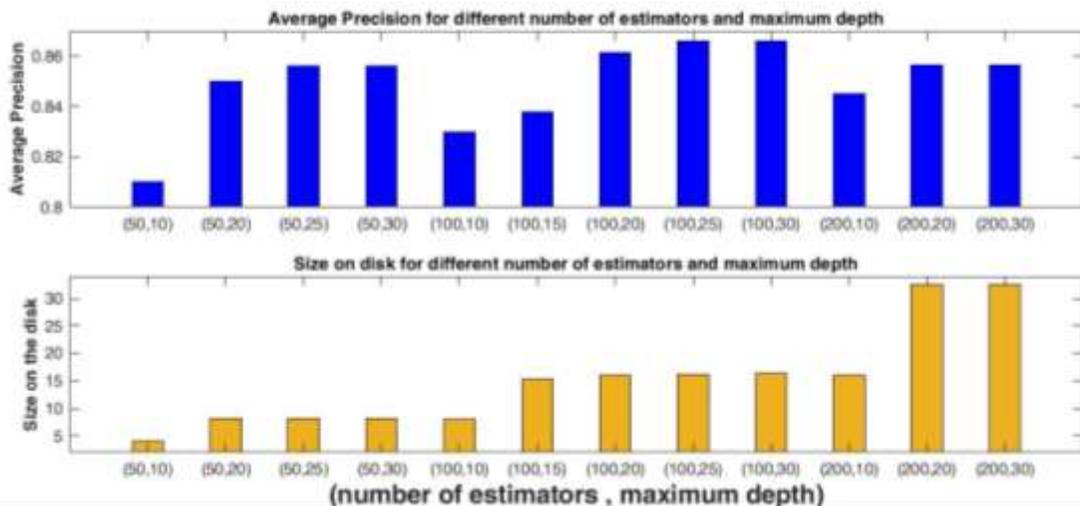


Fig. 3. Average Precision and Size on Disk for Several Combinations of the Number of Trees (Estimators) and the Maximum Depth in the RF.

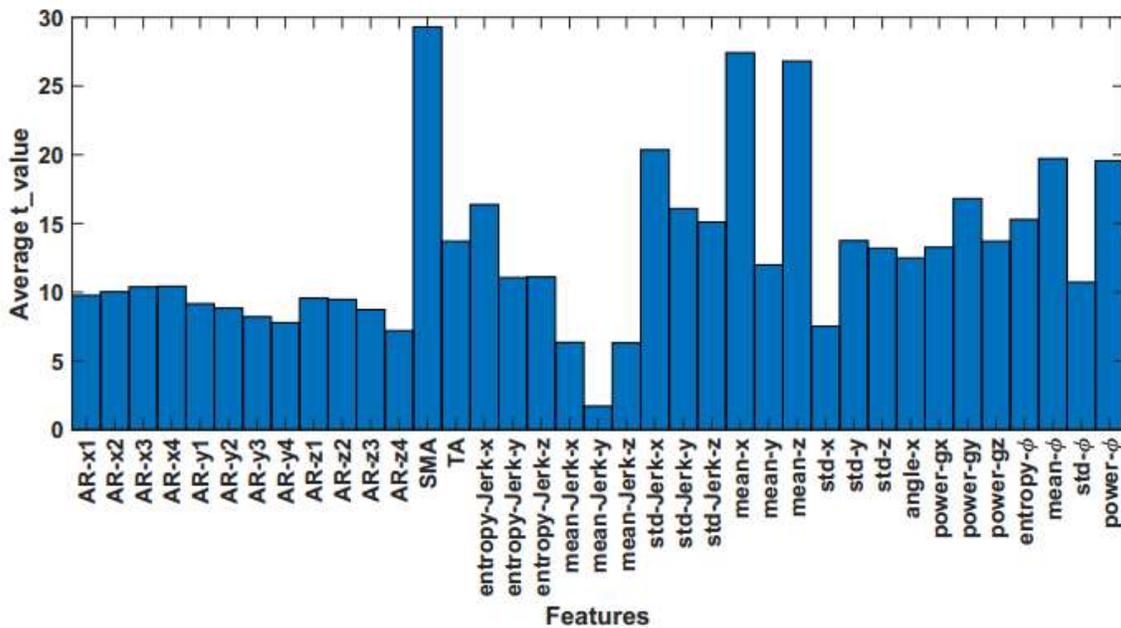


Fig. 4. Estimated Discriminating Power of Features using T-Test.

TABLE III. PRECISION (P) FOR PROPOSED MODEL USING FULL SET OF 37 FEATURES, FEATURES WITH NO SCALING AND REDUCED SET OF 24 FEATURES FOR EACH ACTIVITY CLASS IN WHARF DATASET FOR RANDOM FOREST WITH 100 ESTIMATORS AND DEPTH 20

| Feature Settings | BT | CS | CH | DS | DG | GB | LB | PW | SD | SU | UT | WK | Av. Pre. (%) |
|-------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|--------------|
| Model feature set | 94.62 | 85 | 91.04 | 94.12 | 90.76 | 75.17 | 72.22 | 81.62 | 88.89 | 85.11 | 92.68 | 82.39 | 86.1 |
| No scaling | 89.79 | 85.58 | 92.75 | 94.59 | 79.26 | 74.81 | 80 | 72.87 | 78 | 80 | 85.36 | 89.24 | 83.5 |
| Reduced Features | 94.38 | 86.27 | 87.5 | 89.74 | 83.19 | 62.89 | 100 | 76.69 | 66.67 | 64.06 | 88.09 | 87.03 | 82.2 |

TABLE IV. SIZE ON DISK, TRAINING AND INFERENCE TIME FOR PROPOSED RF, ANN AND SVM CLASSIFIERS USING FULL 37 SET OF NORMALIZED FEATURES

| | Size on the disk | Training time (seconds) | Inference time (seconds) |
|-------------------------------|------------------|--|--------------------------|
| RF (100 estimators, depth 20) | 16 MB | 1.85 | .011 |
| ANN (55 hidden neuron) | 113 KB | 18 (1000 iterations) 12 (500 iterations) 2.88 (100 iterations) | .001 |
| SVM (Grid search training) | 930 KB | 150 (via grid search) | < .000001 sec |

E. Comparison with other Studies on WHARF Dataset

In this section, we compare the proposed model to other models on the same dataset [24] and [23]. First, Jordao [24] used the same activities used in the current model. However, the author first performed data augmentation and calculated the attitude estimation as features to improve the convolution neural network performance. On the other hand, Aguirre [23] dealt with the raw data, performed feature extraction and then introduced the features to a SVM classifier. In the aforementioned studies, accuracy the model has been reported, thus for the proposed model here average accuracy is calculated as the ratio of total number of correctly predicted labels to total number of tested labels. Results of classification accuracy shown in Table V reveal the superiority of proposed model.

F. Limitations of the Current Work

The proposed model (features + classifier) is tested only on one dataset. However, in order to well investigate the generalization of such model, there is a need to test more benchmark datasets for human activity recognition. In addition, the model needs to be refined in order to achieve real time requirements such as considering smaller window size. Also, current study lacks to consider the effect of variant sampling rates of employed sensors. For WHARF data set, sensor is 32 Hz whereas sensors embedded in smartphones are usually 50 Hz. Similarly, sensor or device orientation is expected to affect such HAR models performance. Here, the effect of roll angle is considered, however other dynamic movements of human limbs (i.e. wrist, shoulder, waist or leg) are vital for determining the most suitable feature set.

TABLE V. COMPARISON OF THE PERFORMANCE OF THE PROPOSED MODEL TO PREVIOUS STUDIES [24] AND [23] ON THE WHARF DATASET

| | Feature Extraction | Features domain | Classifier | Average Accuracy (%) |
|----------------|--------------------------|-----------------------------------|------------|----------------------|
| Aguirre [23] | Engineered | Time-domain | SVM | 66.48 |
| Jordao [24] | Raw acceleration signals | Activations of convolution layers | CNN | 79.31 |
| Proposed model | Engineered | Time-domain | RF | 84.86 |

VII. CONCLUSION AND FUTURE WORK

In this work, a simple classification model based on random forest classifier has been proposed for human activity recognition tasks. HAR becomes a very attractive field not only due to the wide range of applicability of machine learning tools, but also for important applications like rehabilitation, health monitoring and clinical applications. The proposed technique employs a feature vector consists of several time-domain features extracted from accelerometer sensor data such as AR model coefficients, mean, and standard deviation. The proposed model is shown to achieve better average accuracy compared to other methods proposed in the literature such as SVM and ANN. RF also has a better classification rate compared to CNN on the same WHARF dataset. The proposed system was trained for segmented data as done in some previous studies.

Examining the implementation of the proposed model on smart devices can be examined in future work. It is also possible to use more than one sensor embedded in smartphones instead of using one wearable sensor as in WHARF. This opens a window for an interesting extension in HAR field concerning implementation of efficient and accurate models on personal devices and examining them in practical environments. Another reasonable extension for this work is to deal with signals that may contain readings of more than one activity. For example, the user may be in a continuous movement where he or she switches between some activities like walking, climbing stairs, sitting and others. It is therefore interesting to examine the performance of the proposed models in literature in real-time situations and ensure that they achieve results similar to those obtained off line.

REFERENCES

- [1] X. Liu, L. Liu, S. J. Simske, and J. Liu, "Human daily activity recognition for healthcare using wearable and visual sensing data," in 2016 IEEE International Conference on Healthcare Informatics (ICHI), 2016, pp. 24-31.
- [2] S. Ranasinghe, F. Al Machot, and H. C. Mayr, "A review on applications of activity recognition systems with regard to performance and evaluation," International Journal of Distributed Sensor Networks, vol. 12, p. 1550147716665520, 2016.
- [3] A. Wang, G. Chen, C. Shang, M. Zhang, and L. Liu, "Human activity recognition in a smart home environment with stacked denoising autoencoders," in International conference on web-age information management, 2016, pp. 29-40.
- [4] C. Jobanputra, J. Bavishi, and N. Doshi, "Human Activity Recognition: A Survey," Procedia Computer Science, vol. 155, pp. 698-703, 2019.
- [5] W. Sousa Lima, E. Souto, K. El-Khatib, R. Jalali, and J. Gama, "Human Activity Recognition Using Inertial Sensors in a Smartphone: An Overview," Sensors, vol. 19, p. 3213, 2019.
- [6] W. Qi, H. Su, and A. Aliverti, "A Smartphone-Based Adaptive Recognition and Real-Time Monitoring System for Human Activities," IEEE Transactions on Human-Machine Systems, 2020.
- [7] P. L. Aguirre, L. A. Torres, and A. P. Lemos, "Autoregressive modeling of wrist attitude for feature enrichment in human activity recognition," in Congresso Brasileiro de Inteligência Computacional, 2017.
- [8] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical human activity recognition using wearable sensors," Sensors, vol. 15, pp. 31314-31338, 2015.
- [9] B. Bruno, F. Mastrogiovanni, and A. Sgorbissa, "A public domain dataset for ADL recognition using wrist-placed accelerometers," in the 23rd IEEE International Symposium on Robot and Human Interactive Communication, 2014, pp. 738-743.
- [10] H. Zhang, Z. Xiao, J. Wang, F. Li, and E. Szczerbicki, "A Novel IoT-Perceptive Human Activity Recognition (HAR) Approach Using Multihead Convolutional Attention," IEEE Internet of Things Journal, vol. 7, pp. 1072-1080, 2019.
- [11] S. Oniga and J. Sütő, "Human activity recognition using neural networks," in Proceedings of the 2014 15th International Carpathian Control Conference (ICCC), 2014, pp. 403-406.
- [12] L. B. Marinho, A. H. de Souza Júnior, and P. P. Rebouças Filho, "A new approach to human activity recognition using machine learning techniques," in International Conference on Intelligent Systems Design and Applications, 2016, pp. 529-538.
- [13] R.-A. Voicu, C. Dobre, L. Bajenaru, and R.-I. Ciobanu, "Human physical activity recognition using smartphone sensors," Sensors, vol. 19, p. 458, 2019.
- [14] A. M. Khan, Y.-K. Lee, and T.-S. Kim, "Accelerometer signal-based human activity recognition using augmented autoregressive model coefficients and artificial neural nets," in 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2008, pp. 5172-5175.
- [15] E. De-La-Hoz-Franco, P. Ariza-Colpas, J. M. Quero, and M. Espinilla, "Sensor-based datasets for human activity recognition—a systematic review of literature," IEEE Access, vol. 6, pp. 59192-59210, 2018.
- [16] F. Demrozi, G. Pravadelli, A. Bihorac, and P. Rashidi, "Human Activity Recognition using Inertial, Physiological and Environmental Sensors: a Comprehensive Survey," arXiv preprint arXiv:2004.08821, 2020.
- [17] N. Capela, E. Lemaire, N. Baddour, M. Rudolf, N. Goljar, and H. Burger, "Evaluation of a smartphone human activity recognition application with able-bodied and stroke participants," Journal of neuroengineering and rehabilitation, vol. 13, pp. 1-10, 2016.
- [18] W.-Y. Cheng, A. Scotland, F. Lipsmeier, T. Kilchenmann, L. Jin, J. Schjodt-Eriksen, et al., "Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients," in 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), 2017, pp. 249-250.
- [19] E. Brophy, W. Muehlhausen, A. F. Smeaton, and T. E. Ward, "Optimised Convolutional Neural Networks for Heart Rate Estimation and Human Activity Recognition in Wrist Worn Sensing Applications," arXiv preprint arXiv:2004.00505, 2020.
- [20] A. Subasi, K. Khateeb, T. Brahim, and A. Sarirete, "Human activity recognition using machine learning methods in a smart healthcare environment," in Innovation in Health Informatics, ed: Elsevier, 2020, pp. 123-144.
- [21] M. M. Moussa, E. Hamayed, M. B. Fayek, and H. A. El Nemr, "An enhanced method for human action recognition," Journal of advanced research, vol. 6, pp. 163-169, 2015.
- [22] R. Poppe, "A survey on vision-based human action recognition," Image and vision computing, vol. 28, pp. 976-990, 2010.

- [23] X. Long, B. Yin, and R. M. Aarts, "Single-accelerometer-based daily physical activity classification," in 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009, pp. 6107-6110.
- [24] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, "A public domain dataset for human activity recognition using smartphones," in Esann, 2013.
- [25] M.-W. Lee, A. M. Khan, and T.-S. Kim, "A single tri-axial accelerometer-based real-time personal life log system capable of human activity recognition and exercise information generation," *Personal and Ubiquitous Computing*, vol. 15, pp. 887-898, 2011.
- [26] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer," *IEEE transactions on information technology in biomedicine*, vol. 14, pp. 1166-1172, 2010.
- [27] A. M. Khan, Y.-K. Lee, S.-Y. Lee, and T.-S. Kim, "Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis," in 2010 5th international conference on future information technology, 2010, pp. 1-6.
- [28] G. Krassnig, D. Tautinger, C. Hofmann, T. Wittenberg, and M. Struck, "User-friendly system for recognition of activities with an accelerometer," in 2010 4th International Conference on Pervasive Computing Technologies for Healthcare, 2010, pp. 1-8.
- [29] A. A. Sukor, A. Zakaria, and N. A. Rahim, "Activity recognition using accelerometer sensor and machine learning classifiers," in 2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA), 2018, pp. 233-238.
- [30] W. Hamäläinen, M. Järvinen, P. Martiskainen, and J. Mononen, "Jerk-based feature extraction for robust activity recognition from acceleration data," in 2011 11th International Conference on Intelligent Systems Design and Applications, 2011, pp. 831-836.
- [31] M. Ghobadi, J. Sosnoff, T. Kesavadas, and E. T. Esfahani, "Using mini minimum jerk model for human activity classification in home-based monitoring," in 2015 IEEE International Conference on Rehabilitation Robotics (ICORR), 2015, pp. 909-912.
- [32] A. Jordao, L. A. B. Torres, and W. R. Schwartz, "Novel approaches to human activity recognition based on accelerometer data," *Signal, Image and Video Processing*, vol. 12, pp. 1387-1394, 2018.
- [33] M. Altuve, P. Lizarazo, and J. Villamizar, "Human activity recognition using improved complete ensemble EMD with adaptive noise and long short-term memory neural networks," *Biocybernetics and Biomedical Engineering*, 2020.
- [34] K. Xia, J. Huang, and H. Wang, "LSTM-CNN Architecture for Human Activity Recognition," *IEEE Access*, vol. 8, pp. 56855-56866, 2020.
- [35] C. A. Ronao and S.-B. Cho, "Human activity recognition with smartphone sensors using deep learning neural networks," *Expert systems with applications*, vol. 59, pp. 235-244, 2016.
- [36] E. Kim, "Interpretable and Accurate Convolutional Neural Networks for Human Activity Recognition," *IEEE Transactions on Industrial Informatics*, 2020.
- [37] P. L. R. Aguirre, "Feature Enrichment in Human Activity Recognition."