

## Research Article

# Effect of Data Segmentation on the Quality of Human Activity Recognition

Basma A. Atalaa<sup>1</sup>, Ahmed Alenany<sup>2</sup>, Ahmed Helmi<sup>2</sup> and Ibrahim Ziedan<sup>3</sup><sup>1</sup>Master student at Computer Engineering and Systems Department, Faculty of Engineering, Zagazig University, Egypt<sup>2</sup>Lecturer at Computer Engineering and Systems Department, Faculty of Engineering, Zagazig University, Egypt<sup>3</sup>Professor at Computer Engineering and Systems Department, Faculty of Engineering, Zagazig University, Egypt**Article History**

Received: 12.06.2020

Accepted: 29.06.2020

Published: 06.07.2020

**Journal homepage:**<https://www.easpublisher.com/easjecs>**Quick Response Code**

**Abstract:** In this paper, we introduce an Artificial Neural Network (ANN) classifier for human activity recognition. The proposed system is divided to three stages; first we do segmentation for raw data collected for two data sets WHARF and UCI-HAR to segment length 128 and 256 with 50% overlap for WHARF, and 128, 256 and 512 with 50% overlap for UCI\_HAR. Second, for each segment, a set of time and frequency domain features are extracted and delivered to the ANN classifier. From a practical point of view, activity classification based on segments of data compared to the use of whole raw data is more suitable and enables faster classification process, especially for short activities. The proposed system is tested against other classifiers such as support vector machines (SVM), naive Bayes, and  $k$ -nearest neighbor (KNN) where ANN gives the best recognition rate. For WHARF dataset, the average accuracy is 68% for segment length 128 and 80% for 256 segment length. On the other hand, employing accelerometer data only in UCI-HAR dataset, the average accuracy is 93.9% for segment length 128, 94.6% for segment length 256 and 96.3% for segment length 512. While using gyroscope data in the same dataset results in an average accuracy of 77.96% for segment length 128, 83.75% for segment length 256 and 88.96% for segment length 512.

**Keywords:** Human Activity Recognition, Artificial Neural Networks, Data Segmentation, Segment length, Training Time

**Copyright © 2020 The Author(s):** This is an open-access article distributed under the terms of the Creative Commons Attribution **4.0 International License (CC BY-NC 4.0)** which permits unrestricted use, distribution, and reproduction in any medium for non-commercial use provided the original author and source are credited.

## INTRODUCTION

Human Activity Recognition (HAR), which is the process of recognizing daily activities (DA) of people, became an active research topic in recent days. HAR can be used to monitoring and taking care of elderly and disabled people as well as children (Planinc, R., & Kampel, M. 2013, December). It is also used in home automation e.g. automatic monitoring of lighting, computers, heating and air conditions (Al Zamil, M. G. *et al.*, 2017).

HAR can be divided into two approaches, The first is based on visual data (computer vision) (Moussa, M. M. *et al.*, 2015; & Poppe, R. 2010) and the second employs data collected from one or more sensors such as tri-axial accelerometer (Aguirre, P. L. *et al.*, 2017; Aguirre, P. L. *et al.*, 2015; & Bruno, B. *et al.*, 2014, August), tri-axial Gyroscope (Marinho, L. B. *et al.*, 2016, December; & Voicu, R. A. *et al.*, 2019), magnetometer (Voicu, R. A. *et al.*, 2019) or heart rate belt sensor with ordinary sensor (Oniga, S., & Sütő, J. 2014, May). The latter approach is more usable due to the availability of low cost sensors especially those available in smart phones (Jobanputra, C. *et al.*, 2019;

Sousa Lima, W. *et al.*, 2019; & Shoaib, M. *et al.*, 2016) and the low computational resources required compared to image processing based approach.

Activity Recognition (AR) employing sensory data can be implemented on raw data (Aguirre, P. L. *et al.*, 2017) or segmented data (Planinc, R., & Kampel, M. 2013, December; Al Zamil, M. G. *et al.*, 2017; & Anguita, D. *et al.*, 2012, December). The use of segmented data is more suitable and enables faster classification process, especially for short activities (Nadi, R. A., & Zamil, M. G. A. 2019; Bulling, A. *et al.*, 2014; Hammerla, N. Y. *et al.*, 2016; Yala, N. *et al.*, 2015, September; Li, K. *et al.*, 2019; & Banos, O. *et al.*, 2014). Data segmentation can be done by dividing the data stream into segments fixed size (Time-based windowing). This is the most used segmentation technique as it is easy to implement. However, this technique requires a careful choice of the segment length. A short segment may be insufficient for activity recognition, while a long segment may contain data belonging to more than one activity leading to poor classification accuracy. In the other hand there is another type for data segmentation called Sensor-based

windowing in which data is divided into windows of equal number of sensor events (Yala, N. *et al.*, 2015, September).

The process of human activity recognition can be divided into two stages. The first stage is feature extraction and used to reduce data dimensionality for raw or segmented data and the second is the classification stage. A preprocessing stage, before feature extraction, may be required for data filtering from noise ( Bruno, B. *et al.*, 2014, August) or data segmentation (Anguita, D. *et al.*, 2012, December).

The features extracted from raw data are diverse. However, they can be broadly classified into time and frequency domain features such as: First, time domain features: minimum, maximum (Sukor, A. A. *et al.*, 2018, March; & Anguita, D. *et al.*, 2013, April), mean (Voicu, R. A. *et al.*, 2019; & Long, X. *et al.*, 2009, September), standard deviation (Aguirre, P. L. *et al.*, 2015; Oniga, S., & Sütő, J. 2014, May; Anguita, D. *et al.*, 2013, April; & Long, X. *et al.*, 2009, September), median (Aguirre, P. L. *et al.*, 2015; & Anguita, D. *et al.*, 2013, April), signal magnitude area (SMA) (Aguirre, P. L. *et al.*, 2017; Anguita, D. *et al.*, 2013, April; Khan, A. M. *et al.*, 2008, August; Lee, M. W. *et al.*, 2011; Khan, A. M. *et al.*, 2010; & Krassnig, G. *et al.*, 2010 March), tilt angle (Aguirre, P. L. *et al.*, 2017; Anguita, D. *et al.*, 2012, December; Sukor, A. A. *et al.*, 2018, March; Anguita, D. *et al.*, 2013, April; Khan, A. M. *et al.*, 2008, August; Lee, M. W. *et al.*, 2011; Khan, A. M. *et al.*, 2010; & Krassnig, G. *et al.*, 2010 March), average absolute difference, average resultant acceleration and histogram (Voicu, R. A. *et al.*, 2019), skewness, kurtosis (Aguirre, P. L. *et al.*, 2015), correlation between axes, energy and spectral entropy (Anguita, D. *et al.*, 2013, April; & Krassnig, G. *et al.*, 2010, March), coefficients of autoregressive model for 3-axis  $x$ ,  $y$  and  $z$  (Aguirre, P. L. *et al.*, 2017; Anguita, D. *et al.*, 2013, April; Khan, A. M. *et al.*, 2008, August; Lee, M. W. *et al.*, 2011; Khan, A. M. *et al.*, 2010; Khan, A. M. *et al.*, 2010, May) and tilt angle (Aguirre, P. L. *et al.*, 2017; & Khan, A. M. *et al.*, 2008, August). Second, frequency domain features : discrete Fourier transform and dc component (Aguirre, P. L. *et al.*, 2015), frequency signal kurtosis (Anguita, D. *et al.*, 2013, April), and power spectral density (PSD) (Aguirre, P. L. *et al.*, 2015; & Nadi, R. A., & Zamil, M. G. A. 2019).

The employment of large number of features is very important because each class has its own discriminating features that can be used to distinguish it from other classes. For example, standard deviation is used to distinguish between static and dynamic classes and Fast Fourier Transform can be used to distinguish between walking and running (Oniga, S., & Sütő, J. 2014, May). To detect the most dominant features and reduce the number of features extracted Principle component analysis (PCA) algorithm (Long, X. *et al.*, 2009, September).

After feature extraction, the activity recognition is performed with the aid of a suitable classifier. In the literature, there are many types of classifiers used in HAR such as neural networks (Voicu, R. A. *et al.*, 2019; Sukor, A. A. *et al.*, 2018, March; Khan, A. M. *et al.*, 2008, August; & Lee, M. W. *et al.*, 2011; Krassnig, G. *et al.*, 2010, March; & Khan, A. M. *et al.*, 2010, May), Support Vector Machine (SVM) (Aguirre, P. L. *et al.*, 2017; Aguirre, P. L. *et al.*, 2015; Marinho, L. B. *et al.*, 2016, December; Sukor, A. A. *et al.*, 2018, March; Anguita, D. *et al.*, 2013, April; & He, Z. Y. *et al.*, 2008, July), decision tree (Sukor, A. A. *et al.*, 2018, March; & Krassnig, G. *et al.*, 2010 March), kernel discriminant analysis (Khan, A. M. *et al.*, 2010, May), Linear discriminant analysis (Lee, M. W. *et al.*, 2011; & Khan, A. M. *et al.*, 2010), random forest (Aguirre, P. L. *et al.*, 2015), Gaussian mixture model (GMM) (Aguirre, P. L. *et al.*, 2015; & Bruno, B. *et al.*, 2014, August), naïve Bayes classifier (Marinho, L. B. *et al.*, 2016, December; & Long, X. *et al.*, 2009, September),  $k$ -Nearest Neighbors (kNN) (Aguirre, P. L. *et al.*, 2015; & Marinho, L. B. *et al.*, 2016, December),  $k$ -means clustering and hidden Markov model (HMM) (Aguirre, P. L. *et al.*, 2015) and minimal learning machine (MLM) (Marinho, L. B. *et al.*, 2016, December).

In this paper, a feed forward neural network (FFN), with one hidden layer, is used and a set of time domain features, namely the coefficients of an autoregressive model, tilt angle, signal magnitude area, Histogram data (Voicu, R. A. *et al.*, 2019), mean, standard deviation, kurtosis and skewness. The number of neurons in the hidden layer for the FFN, the order of AR model and number of histogram bins are selected according to the statistical  $t$ -test in order to achieve the best results. Instead of using raw data only we apply data segmentation to two datasets WHARF and UCI-HAR with segment length of 128, 256, and 512 samples and 50% overlap.

#### Time domain and statistical features

In this section, we review the features to be extracted from the segmented data. In both WHARF and UCI-HAR datasets, there are 3-axis data  $x$ ,  $y$  and  $z$  for the accelerometer  $ax(i)$ ,  $ay(i)$  and  $az(i)$  and gyroscope  $gx(i)$ ,  $gy(i)$  and  $gz(i)$  where  $i=1,2,3,\dots,N$ , and  $N$  is the number of samples in the segment. The following features are used:

**Autoregressive (AR) model:** Autoregressive model is an empirical model used to describe or predict time series data from past data records. The autoregressive model takes the following form for the sensor readings in the  $x$ -direction

$$ax(n) = - \sum_{k=1}^p a(k) ax(n-k) + e(n), \quad (1)$$

where  $ax(n)$  is the signal at time  $n$ ,  $ax(n - 1)$ , ..... ,  $ax(n - p)$  are the values of the signal at past time lags where the maximum lag  $p$  denotes the order of the model. The coefficients  $a(k)$  are the model parameters to be estimated and used for predicting future signal values. The input  $e(n)$  denotes model residual or error which is zero-mean white noise sequence. For the purpose of the current work, the autorgeressive model (1) is applied to the segmented data in each of the  $x$ ,  $y$ , and  $z$  directions, forming a total of  $3p$  features. For model order 3, we have 9 features.

**Signal Magnitude Area (SMA):** is a scalar feature used to distinguish between static and dynamic activities such as standing and walking (Aguirre, P. L. et al., 2017). It is calculated as

$$SMA = \left(\frac{1}{N}\right) (\sum_{i=1}^N |ax(i)| + |ay(i)| + |az(i)|) \quad (2)$$

where  $N$  is the number of samples, and  $ax(i)$ ,  $ay(i)$  and  $az(i)$  are the acceleration in  $x$ ,  $y$  and  $z$  directions, respectively.

**Tilt angle:** is the angle between the  $z$ -axis and the gravitational vector  $g$  and is used to distinguish postures such as standing and lying (Aguirre, P. L. et al., 2017). It is expressed as

$$\Phi = \frac{1}{N} \sum_{i=1}^N \arcsin\left(\frac{az(i)}{\|az\|}\right) \quad (3)$$

Where  $\|az\|$  denotes the 2-norm of the whole accelerometer readings in the  $z$  direction  $az$ .

**Histogram:** A histogram provides an empirical estimate of the probability density function (PDF) of a given random variable from a series of measurements. The range of measurements is first split into  $L$  disjoint bins. The number of measurements  $hi$  falling in bin  $i$  is normalized by the total number of measurements  $N$  to yield the histogram  $\{h1, h 2, \dots, hL\}$ . As the histogram is applied to the data in  $x$ ,  $y$ , and  $z$  directions, this

constitutes a total of  $3L$  features. We split the data into 5 disjoint bins so we have 15 features.

**Mean:** The mean describes the central tendency or the dc level of the data and is calculated as:

$$\mu_x = \frac{1}{N} \sum_{i=1}^N ax(i) \quad (4)$$

**Standard deviation:** describes the amount of variation around the mean and is calculated as:

$$\sigma = \frac{1}{N} \sum_{i=1}^N (ax(i) - \mu_x)^2 \quad (5)$$

**Skewness:** expresses the degree of (a) symmetry of the probability density function of the distribution generating the data and is found as:

$$s = \frac{\frac{1}{N} \sum_{i=1}^N (ax(i) - \mu_x)^3}{\left(\frac{1}{N-1} \sum_{i=1}^N (ax(i) - \mu_x)^2\right)^{3/2}} \quad (6)$$

**Kurtosis:** the kurtosis denotes the degree of tailed-ness of the density function (Helwig, N. et al., 2015, May) and is expressed as:

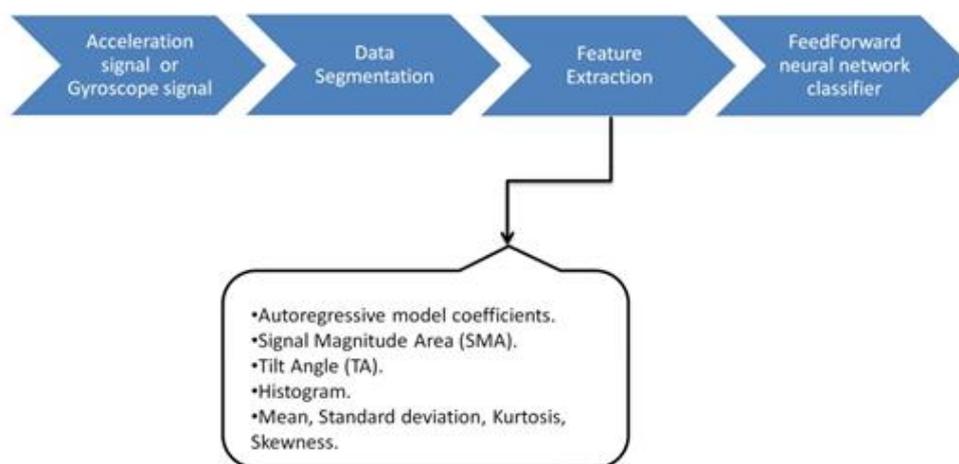
$$k = \frac{\frac{1}{N} \sum_{i=1}^N (ax(i) - \mu_x)^4}{\left(\frac{1}{N} \sum_{i=1}^N (ax(i) - \mu_x)^2\right)^2} - 3 \quad (7)$$

Several previous studies have considered these four statistics (mean, standard deviation, skewness and kurtosis), however, in other applications (Helwig, N. et al., 2015, May; & Chawathe, S. S. 2019, January).

**The proposed classifier**

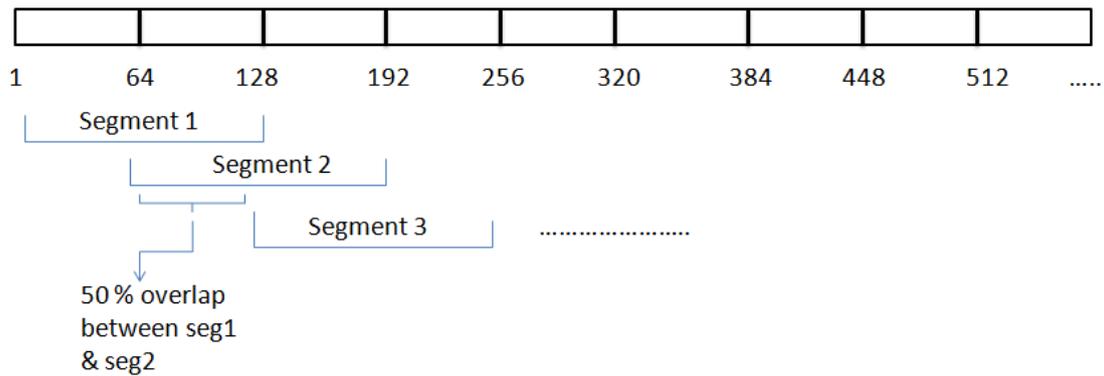
The proposed classifier consists of three stages as shown in Fig 1. The first stage performs data segmentation. The aforementioned features are extracted from raw sensor data in the second stage and the classification task is done in the third stage using feedforward neural network (FNN) (Russell, S. J., & Norvig, P. 2010).

In this work we use segmentation with fixed segment length and 50% overlap between segments as shown in Fig. 2.



**Figure 1:** Block diagram of the proposed human activity recognition system.

Segmentation with segment length 128 and 50% overlap



**Figure 2:** Data Segmentation with segments of length 128 and 50% overlap.

The structure of a typical FNN network consists of the input, one hidden, and the output layers. The input layer receives a set of  $ns$  ( $3p + 3L + 14$ ) features where  $p$  is the order of the autoregressive model,  $L$  is the number of histogram bins, and  $ns$  is the number of sensors employed. So, in this work, we have a total of 38 features. Both the hidden and output layers consist of units called neurons employing a linear combination of the outputs of the previous layer followed by a nonlinearity function which is typically the sigmoid function defined as

$$y = \frac{1}{1 + e^{-x}} \tag{8}$$

Where  $x$  is the linear combination sum and  $y$  is the nonlinear neuron output. While the output layer consists of  $m$  neurons, where  $m$  denotes the number of activity classes.

The features presented to the neural networks must be normalized so that all features lie in the same range, e.g.,  $(-1, 1)$ . On the other hand, the outputs of the neurons in the output layer are real numbers and so are not directly suitable for classification. For this purpose, the output  $y_i$  of each neuron is transformed using the following so-called softmax function

$$c_i = \frac{e^{y_i}}{\sum_{i=1}^m e^{y_i}} \tag{9}$$

Where  $c_i$  is constrained in the interval  $(0, 1)$  and all  $c_i, i = 1, 2, \dots, m$  sums to 1.  $c_i$  can therefore be interpreted as a probability and so the classification decision can be made as the class corresponding to the maximum  $c_i$ . Training neural networks involves updating their weights until they achieve a minimum cost function. The default cost function is the mean square error (MSE) calculated as

$$mse = \frac{1}{Nt} \frac{1}{m} \sum_{i=1}^{Nt} \sum_{k=1}^m (tk(i) - ck(i))^2 \tag{10}$$

Where  $tk(i)$  is  $k$ -th target output and  $ck(i)$  is the  $k$ -th network output at the  $i$ -th training example,  $Nt$  is the number of training examples presented to the network, and  $m$  is the number of classes. The training of FNN network is performed using the back-propagation algorithm.

## RESULTS AND DISCUSSION

### Datasets

In this section, two benchmark HAR datasets, namely WHARF and UCI-HAR, are used to examine the performance of the proposed technique. The datasets differ according to (1) considered activity signals, (2) used sensor, (3) sensor position on human body and (4) experimental setting for recording activity signals. The following are brief of both datasets.

### WHARF

Wearable Human Activity Recognition Folder (WHARF) dataset was prepared by Bruno *et al.*, (2014, August). The dataset was collected by an ad-hoc tri-axial accelerometer sensor attached to right wrist of each of 17 participants; 11 males, with age ranging from 19 to 81 years; and 6 females, with ages between 56 and 85 years (Aguirre, P. L. *et al.*, 2017). Experiments were taking place at each participant home in a supervised manner. The digital resolution of the sensor was 6 bits and sampling rate was 32 Hz. The dataset contains 12 activities, of which the following 7 activity classes are represented by more than 100 instances and will be considered in this paper for the reasonable training of the proposed model: *Climb stairs* (CS), *Drink glass* (DG), *Getup bed* (GB), *Pour water* (PW), *Sitdown chair* (SD), *Standup chair* (SU) and *Walk* (WK). Each activity class examples are contained in a separate folder and raw signals for each single activity are saved in one text file.

**UCI-HAR**

A public dataset for activities of daily living have been published by Anguita *et al.*, (2013, April). Participants were asked to follow some protocol to act 6 activities using a waist-mounted smartphone. The activities are namely *Laying* (LY), *Sitting* (ST), *Standing* (SD), *Walking* (WK), *Walking upstairs* (WU) and *Walking down* (WD). A sampling rate of 50 Hz was used to collect the triaxial linear acceleration and angular velocity of the phone accelerometer and gyroscope sensors. Each participant was performing a sequence of activities in order. Hence, raw signals of all activities were recorded in one text file per participant. After parsing the files of raw signals, we can collect the samples of each activity in order to train our model. From each activity class, a number of 120 examples were randomly selected. The number of instances for each activity class is the same for both accelerometer and gyroscope experiments. In a previous published version of this dataset, only data for some extracted features of the activities were available. Hence, most studies in the literature considering this dataset employed features data rather than the raw signal data.

**Experimental Results**

In this section, the results of several experiments are presented. First, performance of

proposed classification model, in terms of True Positive Rate (TPR) or the classifier recall, accuracy and confusion matrix, is shown for both datasets considered here. Also performance of proposed model is compared to previous studies for both datasets. Second, our ANN classifier is compared with four commonly used classifiers in literature, namely Support Vector Machines, Naive Bayes, and *K*-Nearest Neighbors. The average overall accuracy and standard deviation for all classifiers are reported. Finally, the results of proposed model are shown for activity classes in UCI-HAR dataset when combining features of both accelerometer and gyroscope sensors. Experiments were carried out using Matlab on a CPU i5, 2.6 GHz and 6GB RAM. In each experiment, 100 independent computer runs are performed. For the purpose of training the proposed ANN classifier, each dataset has been partitioned randomly into 85% for training and validation, and 15% for testing.

For WHARF dataset, seven activity classes with number of examples greater than 100, namely *Climb stairs*, *Drink glass*, *Getup bed*, *Pour water*, *Sitdown chair*, *Standup chair* and *Walk*, are considered for data segmentation. For balanced training preferred for ANN or other classifiers, we deal with specific number of segments according to the segment length.

**Table 1:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in WHARF dataset segmented with segment length 128 with 50% overlap. The activities are Climb stairs (CS), Drinking glass (DG), Getup bed (GB), Pour water (PW), Sitdown chair (SD), Standup chair(SU) and Walk (WK). Rows are predicted classes while columns are true classes.

	CS	DG	GB	PW	SD	SU	WK
CS	153	3	10	2	5	7	57
DG	1	195	1	30	6	5	0
GB	9	5	141	10	31	35	6
PW	0	21	4	205	5	3	0
SD	5	8	28	10	132	49	5
SU	7	6	25	11	42	138	10
WK	50	0	7	0	6	6	170
P	.7	.8	.7	.8	.6	.6	.7
R	.6	.8	.6	.9	.6	.6	.7
F	.6	.8	.7	.8	.6	.6	.7

**Table 2:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in WHARF dataset segmented with segment length 256 with 50% overlap. Activities labels are the same as in Table 1. Rows are predicted classes while columns are true classes.

	CS	DG	GB	PW	SD	SU	WK
CS	36	0	3	0	1	0	8
DG	0	46	0	2	0	0	0
GB	1	0	37	1	4	4	1
PW	0	3	0	44	1	0	0
SD	1	0	4	1	33	8	0
SU	1	0	3	0	8	36	1
WK	9	0	2	0	0	0	38
P	.75	.96	.77	.92	.7	.73	.77
R	.75	.94	.75	.94	.7	.75	.79
F	.75	.95	.76	.93	.7	.74	.78

**Table 3:** Confusion matrix, Recall (**R**) and **F** measure for proposed model for each activity class in WHARF dataset for raw data. Activities labels are the same as in Table 1. Rows are predicted classes while columns are true classes.

	<b>CS</b>	<b>DG</b>	<b>GB</b>	<b>PW</b>	<b>SD</b>	<b>SU</b>	<b>WK</b>
<b>CS</b>	90	0	1	0	1	1	5
<b>DG</b>	0	93	0	6	0	0	0
<b>GB</b>	3	0	93	0	3	3	1
<b>PW</b>	1	7	0	94	0	1	0
<b>SD</b>	0	0	3	0	87	8	0
<b>SU</b>	2	0	3	0	9	87	1
<b>WK</b>	4	0	0	0	0	0	93
<b>P</b>	.91	.93	.93	.91	.88	.85	.95
<b>R</b>	.9	.93	.93	.94	.87	.87	.93
<b>F</b>	.9	.93	.93	.92	.87	.85	.93

As segment length increases, the results increase as well. A segment length of 256 gives the best

accuracy against 128 segment length and raw data give the best accuracy for WHARF data set.

**Table 4:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR accelerometer dataset segmented with segment length 128 with 50% overlap. The activities are Laying (LY), Sitting (ST), Standing (SD), Walking (WK), Walking upstairs (WD) and Walking down (WU). Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	900	0	0	0	0	0
<b>ST</b>	0	777	121	1	0	1
<b>SD</b>	0	91	808	1	0	0
<b>WK</b>	0	0	0	868	17	15
<b>WD</b>	0	0	0	25	858	17
<b>WU</b>	0	0	0	28	13	859
<b>P</b>	1	.89	.87	.94	.96	.96
<b>R</b>	1	.86	.89	.96	.95	.95
<b>F</b>	1	.87	.87	.95	.95	.95

**Table 5:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR accelerometer dataset segmented with segment length 256 with 50% overlap. Activities labels are the same as in Table 4. Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	354	0	0	0	0	0
<b>ST</b>	0	311	42	0	0	1
<b>SD</b>	0	28	326	0	0	0
<b>WK</b>	0	0	0	342	3	9
<b>WD</b>	0	0	0	6	343	5
<b>WU</b>	0	0	0	16	4	334
<b>P</b>	1	.91	.88	.94	.98	.95
<b>R</b>	1	.87	.92	.96	.96	.94
<b>F</b>	1	.88	.89	.94	.96	.94

**Table 6:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR accelerometer dataset segmented with segment length 512 with 50% overlap. Activities labels are the same as in Table 4. Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	104	0	0	0	0	0
<b>ST</b>	0	96	8	0	0	0
<b>SD</b>	0	4	100	0	0	0
<b>WK</b>	0	0	0	101	0	2
<b>WD</b>	0	0	0	2	102	1
<b>WU</b>	0	0	0	4	1	99
<b>P</b>	1	.96	.92	.94	.99	.97
<b>R</b>	1	.92	.96	.98	.97	.95
<b>F</b>	1	.93	.93	.95	.98	.96

**Table 7:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR accelerometer dataset for raw data. The activities are Laying (LY), Sitting (ST), Standing (SD), Walking (WK), Walking upstairs (WD) and Walking down (WU). Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	120	0	0	0	0	0
<b>ST</b>	0	108	10	0	0	0
<b>SD</b>	0	12	110	0	0	0
<b>WK</b>	0	0	0	107	2	9
<b>WD</b>	0	0	0	3	115	2
<b>WU</b>	0	0	0	10	3	109
<b>P</b>	1	.91	.9	.91	.95	.89
<b>R</b>	1	.9	.91	.89	.95	.9
<b>F</b>	1	.9	.9	.89	.95	.89

It is noted that for the walking activity in UCI-HAR accelerometer the recall 1 which means that it is detected correctly in all cases. As segment length increase the accuracy increase and gives the best accuracy at segment length 512 which also gives more

better accuracy than using raw data for UCI-HAR Accelerometer but the difference in the recall for raw data and segmented data with segment length 512 is not large.

**Table 8:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR gyroscope dataset segmented with segment length 128 and 50% overlap. The activities are Laying (LY), Sitting (ST), Standing (SD), Walking (WK), Walking upstairs (WD) and Walking down (WU). Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	588	159	149	1	2	0
<b>ST</b>	169	625	101	1	0	4
<b>SD</b>	123	135	637	2	1	1
<b>WK</b>	2	0	0	790	65	43
<b>WD</b>	1	0	0	120	759	20
<b>WU</b>	0	0	0	75	13	811
<b>P</b>	.66	.68	.71	.79	.90	.92
<b>R</b>	.65	.69	.7	.87	.84	.9
<b>F</b>	.65	.68	.7	.82	.86	.9

**Table 9:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR gyroscope dataset segmented with segment length 256 and 50% overlap. Activities labels are the same as in Table 8. Rows are predicted classes while columns are true classes.

	<b>LY</b>	<b>ST</b>	<b>SD</b>	<b>WK</b>	<b>WD</b>	<b>WU</b>
<b>LY</b>	244	58	51	0	1	0
<b>ST</b>	54	261	38	0	0	0
<b>SD</b>	42	40	271	1	0	1
<b>WK</b>	1	0	0	338	9	6
<b>WD</b>	0	0	0	23	329	2
<b>WU</b>	0	0	0	18	0	336
<b>P</b>	.72	.72	.75	.89	.97	.97
<b>R</b>	.68	.73	.76	.95	.92	.94
<b>F</b>	.69	.72	.75	.91	.94	.95

**Table 10:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR gyroscope dataset segmented with segment length 512 and 50% overlap. Activities labels are the same as in Table 8. Rows are predicted classes while columns are true classes.

	LY	ST	SD	WK	WD	WU
LY	77	17	10	0	0	0
ST	16	80	8	0	0	0
SD	8	9	87	0	0	0
WK	0	0	0	103	1	0
WD	0	0	0	0	104	0
WU	0	0	0	0	0	104
P	.76	.75	.82	1	.99	1
R	.74	.76	.82	1	1	1
F	.74	.75	.82	1	.99	1

**Table 11:** Confusion matrix, Recall (**R**), Precision (**P**) and **F** measure for proposed model for each activity class in UCI-HAR gyroscope for raw data. Activities labels are the same as in Table 8. Rows are predicted classes while columns are true classes.

	LY	ST	SD	WK	WD	WU
LY	93	15	8	0	0	0
ST	17	99	7	0	0	0
SD	10	6	105	0	0	0
WK	0	0	0	117	2	3
WD	0	0	0	2	118	0
WU	0	0	0	2	0	117
P	.8	.8	.86	.95	.98	.98
R	.77	.82	.87	.96	.98	.97
F	.78	.8	.86	.95	.98	.97

As segment length increases, the accuracy increases achieves the best accuracy at segment length 512 for the activity of walking and walking upstairs where the recall is 1. However, for the remaining activities the raw data gives a larger recall than the segmented data.

**Comparison with other classifiers**

In this section we will introduce a comparison between the proposed classifiers using Feedforward Neural Networks and some other classifiers as Support Vector Machine (SVM), naive Bayes, k-nearest

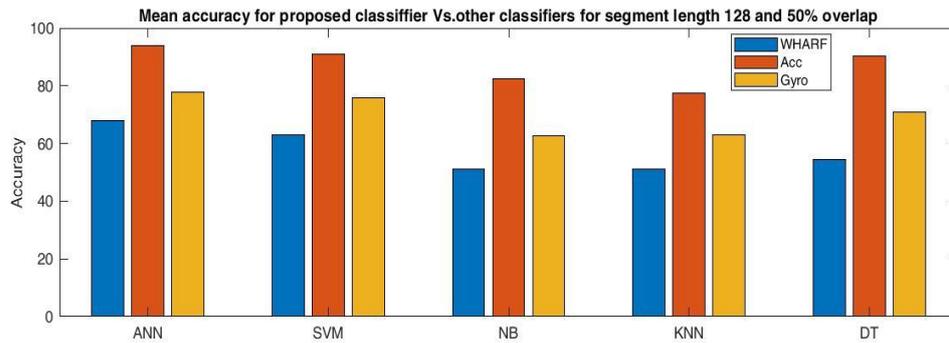
neighbor (KNN) and Decision trees (DT). We compare from two point of view average accuracy and Training time for each algorithm.

**1- Accuracy**

We can see that average accuracy increase as segment length increase for the three datasets, but the best average accuracy for both WHARF and UCI-HAR gyroscope is obtained using raw data. For UCI-HAR accelerometer, the best average accuracy is obtained using the segmented data with segment length 512 with 50% overlap.

**Table 12:** Mean and standard deviation of accuracy (%) of proposed model versus other classifiers for segmented data with segment length 128 and 50% overlap. Average accuracy of artificial neural networks (ANN) is the best for all datasets. **Acc** and **Gyro** are UCI-HAR datasets for accelerometer and gyroscope, respectively. Other classifiers are support vector machines (SVM), Naive Bayes (NB), K-Nearest Neighbors (KNN) and Decision Trees (DT).

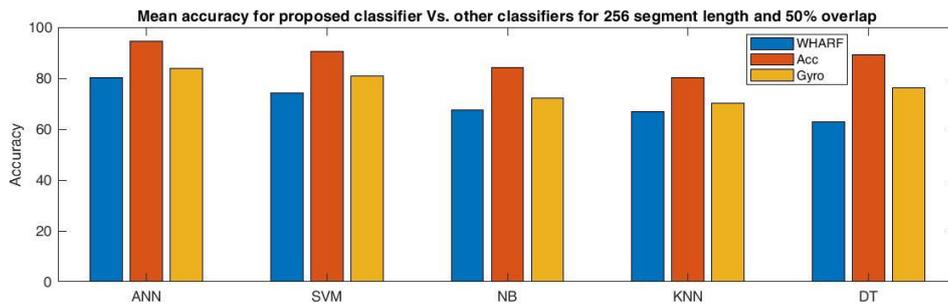
	ANN	SVM	NB	KNN	DT
WHARF	68.03±3.32	63.14±.004	51.31±.0037	51.26±.003	54.39±.009
Acc	93.9±1.11	90.8±.001	82.35±.0009	77.52±.0022	90.45±.0037
Gyro	77.96±2.77	75.7±.0015	62.73±.0019	63.04±.0026	71.04±.0051



**Figure 3:** Mean accuracy for the proposed Artificial Neural Network (ANN) classifier vs. Support Vector Machine (SVM), Naive Bayes (NB), *k*-Nearest Neighbors (KNN) and Decision Trees (DT) for segment length 128 and 50% overlap.

**Table 13:** Mean and standard deviation of accuracy (%) of proposed model versus other classifiers for segmented data with segment length 256 and 50% overlap. Average accuracy of artificial neural networks (ANN) is the best for all datasets. Acc and Gyro are UCI-HAR datasets for accelerometer and gyroscope, respectively. Other classifiers as described in Table 12.

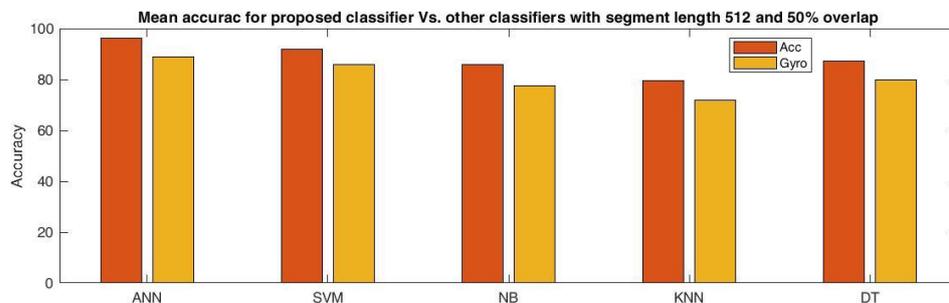
	ANN	SVM	NB	KNN	DT
<b>WHARF</b>	80.11±5.22	74.14±.0104	67.34±.0085	66.81±.0097	62.82±.018
<b>Acc</b>	94.45±1.36	90.62±.0019	84.21±.0021	80.15±.0032	89.08±.0059
<b>Gyro</b>	83.75±2.01	80.74±.0031	72.23±.0064	70.05±.004	76.23±.0066



**Figure 4:** Mean accuracy for proposed classifier vs. other classifiers for segment length 256 and 50% overlap and classifiers as described in Figure 3.

**Table 14:** Mean and standard deviation of accuracy (%) of proposed model versus other classifiers for segmented data with segment length 512 and 50% overlap. Average accuracy of artificial neural networks (ANN) is the best for all datasets. Acc and Gyro are UCI-HAR datasets for accelerometer and gyroscope, respectively. Other classifiers as described in table 12.

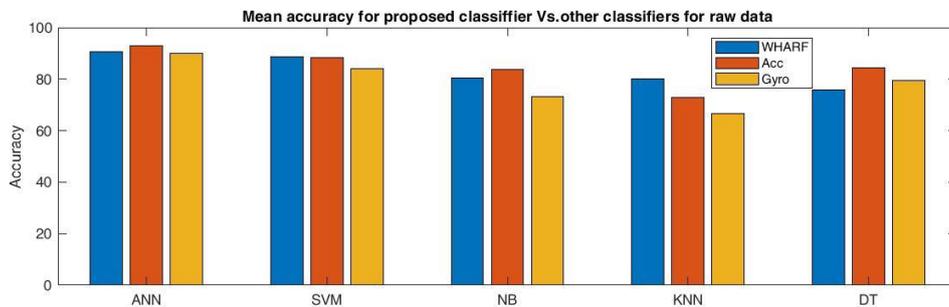
	ANN	SVM	NB	KNN	DT
<b>Acc</b>	96.32±2.62	91.83±.0054	86±.0036	79.36±.0059	87.23±.0086
<b>Gyro</b>	88.96±2.45	85.68±.005	77.61±.0075	71.94±.0073	80±.0106



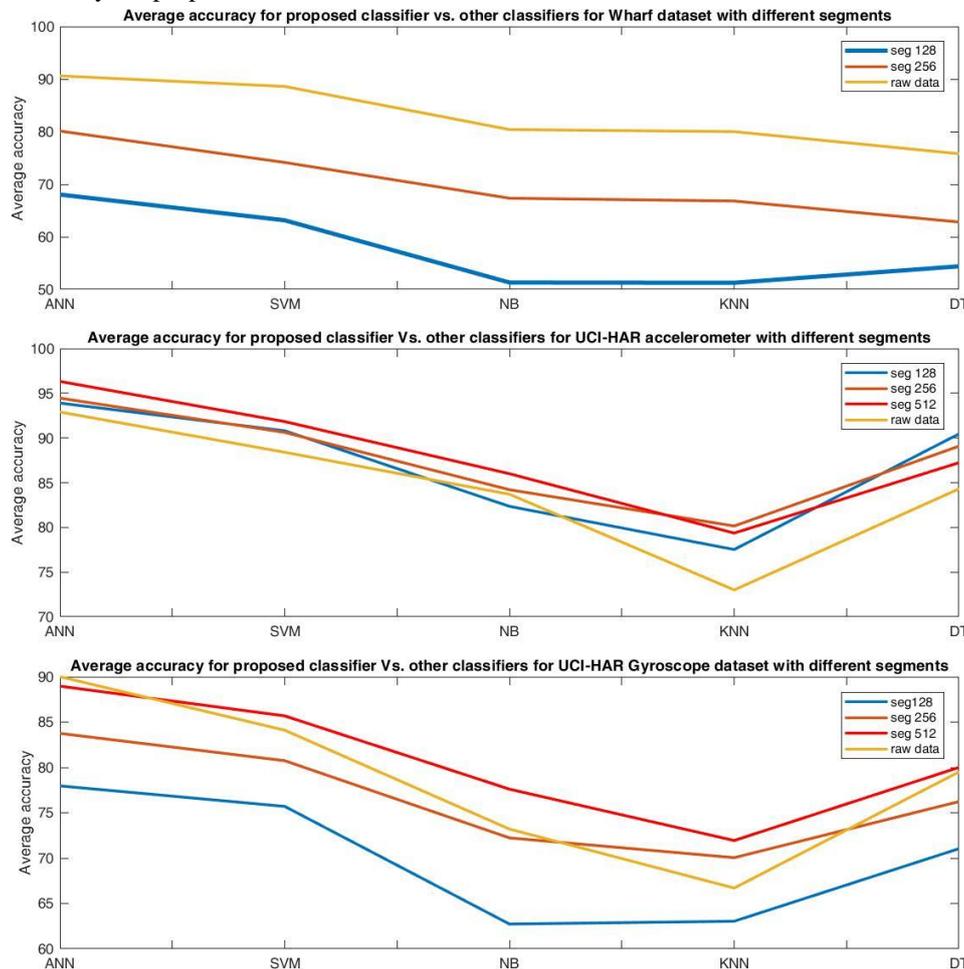
**Figure 5:** Mean accuracy for proposed classifier vs. other classifiers for segment length 512 and 50% overlap and classifiers as described in Figure 3.

**Table 15:** Mean and standard deviation of accuracy (%) of the proposed model versus other classifiers for raw data. Average accuracy of artificial neural networks (ANN) is the best for all datasets. **Acc** and **Gyro** are UCI-HAR datasets for accelerometer and gyroscope, respectively, as described in table 12.

	ANN	SVM	NB	KNN	DT
<b>WHARF</b>	90.6±2.7	88.6±.01	80.4±.01	80±.01	75.8±.02
<b>Acc</b>	92.9±2.7	88.4±.004	83.7±.004	73±.004	84.3±.01
<b>Gyro</b>	90±2	84.1±.003	73.2±.004	66.7±.01	79.5±.01



**Figure 6:** Mean accuracy for proposed classifier vs. other classifiers for raw data and classifiers as described in Figure 3.



**Figure 7 :** Average accuracy for proposed classifier Vs. other classifiers for the used datasets WHARF, UCI-HAR Accelerometer and UCI-HAR Gyroscope with segment lengths 128, 256, 512 and raw data.

**2- Training time**

Here we introduce the comparison of proposed classifier with other classifiers from training time point

of view. Introduce which algorithm give the shortest training time for different segment lengths.

**Table 16:** Training time for various classifiers for the datasets segmented with segment length 128 with 50% overlap.

	#Examples	NN	SVM	NB	KNN	DT
<b>WHARF</b>	238	90.27	6.32*10 <sup>3</sup>	46.52	15.09	132.05
<b>Acc</b>	900	344.59	4.6*10 <sup>3</sup>	58.57	99.19	278.22
<b>Gyro</b>	900	314.19	4.8*10 <sup>3</sup>	58.89	113.4	291.98

**Table 17:** Training time for various classifiers for the datasets segmented with segment length 256 with 50% overlap.

	#Examples	NN	SVM	NB	KNN	DT
<b>WHARF</b>	48	41.12	263.16	41.02	7.75	18.75
<b>Acc</b>	354	126.05	1.08*10 <sup>3</sup>	41.36	27.41	80.88
<b>Gyro</b>	354	123.21	2.35*10 <sup>3</sup>	41.47	23.49	111.11

**Table 18:** Training time for various classifiers for the datasets segmented with segment length 512 with 50% overlap.

	#Examples	NN	SVM	NB	KNN	DT
<b>Acc</b>	104	52.65	317.95	35	8.16	21.6
<b>Gyro</b>	104	48.5	1.07*10 <sup>3</sup>	35.7	8.58	23.25

From Tables 16, 17 and 18, we can see that the training time for the proposed classifier NN is larger than that of NB, KNN and DT. However, the proposed classifier outperforms those classifiers in terms of recognition accuracy. Compared to SVM, the proposed classifier is superior both with regard to the training time and recognition accuracy. Therefore, it can be concluded that the proposed classifier is suitable in real world applications.

The difference in training time between WHARF and UCI-HAR datasets is also acceptable as the number of examples for UCI-HAR dataset is more than in WHARF dataset. For the segments of length 128, we have small training time and faster response compared to large segments. However, in Table 18 we see for segment length 512 that the training time is less than other training time and this can be explained as follows. In our work, we take the same number of examples for each activity so we work for the shortest activity and truncate the rest for large-length activities and this leads to reducing the number of training examples and as a result reducing the training time.

We can see the difference in time for both UCI-HAR Accelerometer and Gyroscope dataset using Neural Network as training algorithm despite using the same number of examples for both. This difference in time results from using different number of neurons in hidden layer. After doing many experiments we found that the best number of neurons for UCI-HAR accelerometer is 90 neuron and, for UCI-HAR gyroscope, it is 75 neuron. As the number of neurons increase the complexity of neural network increase resulting in increasing the time of training the network.

**Conclusions and Future Work**

In this work, a simple classification model based on ANN has been proposed for human activity recognition (HAR) 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 and health monitoring. Two HAR datasets, WHARF and UCI-HAR differing in the types of activity, are considered. The feature vector consists of several time-domain features such as AR model coefficients, histogram values, mean, standard deviations and few others which give best results when applied to ANN classifier.

Performance of ANN was better than other classifiers in machine learning such as SVM, NB, KNN and DT. For WHARF, proposed system was trained for raw signals similar to some previous studies (Aguirre, P. L. et al., 2017; & Aguirre, P. L. R. 2018) and trained for segmented data with segment length 128 and 256 with 50% overlap. For UCI-HAR, proposed system was trained in two cases: one using raw signals and the second using segmented signals with 128 window length and 50% overlapping as suggested by the creators of this dataset(Anguita, D. et al., 2012, December), and also using segment length 256 and 512 with 50% overlap. Average accuracy of the proposed system is the best so far for both examined datasets and the training time also is acceptable for it comparing with other classifiers accuracy.

**REFERENCES**

1. Aguirre, P. L. R. (2018). Feature enrichment in human activity recognition.
2. Aguirre, P. L., Torres, L. A., & Lemos, A. P. (2017). Autoregressive modeling of wrist attitude for feature enrichment in human activity recognition. In *Congresso Brasileiro de Inteligência Computacional*.
3. Al Zamil, M. G., Samarah, S. M., Rawashdeh, M., & Hossain, M. A. (2017). An ODT-based abstraction for mining closed sequential temporal patterns in IoT-cloud smart homes. *Cluster Computing*, 20(2), 1815-1829.

4. Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2012, December). Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine. In *International workshop on ambient assisted living* (pp. 216-223). Springer, Berlin, Heidelberg.
5. Anguita, D., Ghio, A., Oneto, L., Parra, X., & Reyes-Ortiz, J. L. (2013, April). A public domain dataset for human activity recognition using smartphones. In *Esann* (Vol. 3, p. 3).
6. Attal, F., Mohammed, S., Dedabrishvili, M., Chamroukhi, F., Oukhellou, L., & Amirat, Y. (2015). Physical human activity recognition using wearable sensors. *Sensors*, *15*(12), 31314-31338.
7. Banos, O., Galvez, J. M., Damas, M., Pomares, H., & Rojas, I. (2014). Window size impact in human activity recognition. *Sensors*, *14*(4), 6474-6499.
8. Bruno, B., Mastrogiovanni, F., & Sgorbissa, A. (2014, August). A public domain dataset for ADL recognition using wrist-placed accelerometers. In *the 23rd IEEE International Symposium on Robot and Human Interactive Communication* (pp. 738-743). IEEE.
9. Bulling, A., Blanke, U., & Schiele, B. (2014). A tutorial on human activity recognition using body-worn inertial sensors. *ACM Computing Surveys (CSUR)*, *46*(3), 1-33.
10. Chawathe, S. S. (2019, January). Condition Monitoring of Hydraulic Systems by Classifying Sensor Data Streams. In *2019 IEEE 9th Annual Computing and Communication Workshop and Conference (CCWC)* (pp. 0898-0904). IEEE.
11. Hammerla, N. Y., Halloran, S., & Plötz, T. (2016). Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*.
12. He, Z. Y., & Jin, L. W. (2008, July). Activity recognition from acceleration data using AR model representation and SVM. In *2008 international conference on machine learning and cybernetics* (Vol. 4, pp. 2245-2250). IEEE.
13. Helwig, N., Pignanelli, E., & Schütze, A. (2015, May). Condition monitoring of a complex hydraulic system using multivariate statistics. In *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings* (pp. 210-215). IEEE.
14. Jobanputra, C., Bavishi, J., & Doshi, N. (2019). Human activity recognition: a survey. *Procedia Computer Science*, *155*, 698-703.
15. Khan, A. M., Lee, Y. K., & Kim, T. S. (2008, August). 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* (pp. 5172-5175). IEEE.
16. Khan, A. M., Lee, Y. K., Lee, S. Y., & Kim, T. S. (2010). 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.
17. Khan, A. M., Lee, Y. K., Lee, S. Y., & Kim, T. S. (2010, May). Human activity recognition via an accelerometer-enabled-smartphone using kernel discriminant analysis. In *2010 5th international conference on future information technology* (pp. 1-6). IEEE.
18. Krassnig, G., Tantinger, D., Hofmann, C., Wittenberg, T., & Struck, M. (2010, March). User-friendly system for recognition of activities with an accelerometer. In *2010 4th International Conference on Pervasive Computing Technologies for Healthcare* (pp. 1-8). IEEE.
19. Lee, M. W., Khan, A. M., & Kim, T. S. (2011). 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*, *15*(8), 887-898.
20. Li, K., Habre, R., Deng, H., Urman, R., Morrison, J., Gilliland, F. D., ... & Bui, A. A. (2019). Applying multivariate segmentation methods to human activity recognition from wearable sensors' data. *JMIR mHealth and uHealth*, *7*(2), e11201.
21. Long, X., Yin, B., & Aarts, R. M. (2009, September). Single-accelerometer-based daily physical activity classification. In *2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 6107-6110). IEEE.
22. Marinho, L. B., de Souza Júnior, A. H., & Rebouças Filho, P. P. (2016, December). A new approach to human activity recognition using machine learning techniques. In *International Conference on Intelligent Systems Design and Applications* (pp. 529-538). Springer, Cham.
23. Moussa, M. M., Hamayed, E., Fayek, M. B., & El Nemr, H. A. (2015). An enhanced method for human action recognition. *Journal of advanced research*, *6*(2), 163-169.
24. Nadi, R. A., & Zamil, M. G. A. (2019). A profile based data segmentation for in-home activity recognition. *International Journal of Sensor Networks*, *29*(1), 28-37.
25. Oniga, S., & Sütő, J. (2014, May). Human activity recognition using neural networks. In *Proceedings of the 2014 15th International Carpathian Control Conference (ICCC)* (pp. 403-406). IEEE.
26. Planinc, R., & Kampel, M. (2013, December). Detecting changes in elderly's mobility using inactivity profiles. In *International Workshop on Ambient Assisted Living* (pp. 100-103). Springer, Cham.
27. Poppe, R. (2010). A survey on vision-based human action recognition. *Image and vision computing*, *28*(6), 976-990.

28. Russell, S. J., & Norvig, P. (2010). "Artificial Intelligence-A Modern Approach (3rd internat. edn.)," ed: Pearson Education.
29. Shoaib, M., Bosch, S., Incel, O. D., Scholten, H., & Havinga, P. J. (2016). Complex human activity recognition using smartphone and wrist-worn motion sensors. *Sensors*, *16*(4), 426.
30. Sousa Lima, W., Souto, E., El-Khatib, K., Jalali, R., & Gama, J. (2019). Human activity recognition using inertial sensors in a smartphone: An overview. *Sensors*, *19*(14), 3213.
31. Sukor, A. A., Zakaria, A., & Rahim, N. A. (2018, March). Activity recognition using accelerometer sensor and machine learning classifiers. In *2018 IEEE 14th International Colloquium on Signal Processing & Its Applications (CSPA)* (pp. 233-238). IEEE.
32. Voicu, R. A., Dobre, C., Bajenaru, L., & Ciobanu, R. I. (2019). Human physical activity recognition using smartphone sensors. *Sensors*, *19*(3), 458.
33. Yala, N., Fergani, B., & Fleury, A. (2015, September). Feature extraction for human activity recognition on streaming data. In *2015 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA)* (pp. 1-6). IEEE.