Activity Recognizer using Machine Learning Classifiers

Ancy K Biju[1], Ancy P Oommechen[2], Aswathy M T[3], Simi Elizabeth[4]

[1,2,3]B.Tech Student, Computer Science and Engineering, APJ Abdul kalam Technological University, Kerala, India
[4]Asst.Proffessor, Computer Science and Engineering, Mount Zion College of Engineering, Kadammanitta, Kerala, India

---***---

Abstract - *Activity Recognition plays a major role in many applications particularly in day to day life, healthcare services. Medical diagnosis, users daily routine monitoring, abnormal case detection seeks more attention.*

This paper put forwards the approach of using accelerometer sensor embedded in a smartphone which is very much available .The raw input used here is a signal with publically available accelerometer dataset. Time domain and frequency domain feature of the signal are selected here. Principal Component Analaysis. (PCA)used here reduce the dimensionality of features. Among many features, PCA extract most significant features which classify activities of humans. In order to get the accurate results a comparison is performed between original raw data and PCA features. Machine learning approach is provided where the Classifiers compare the time and frequency domain features. The finally obtained result shows high accuracy in frequency domain features and high recognition rate in PCA based features. Comparison technique is highly emphasized here which provides the best results of activity and recognize with high precision.

Key Words **Machine Learning, Principal Component Analysis, Accelerometer data set, Daily living activities, Classification, Comparison**

1. INTRODUCTION

Activity Recognition plays a major role in healthcare services and now a days studied as a part to reduce the cost and workload of officials. It mainly focuses on healthcare services especially on mental health of people, physical illness etc, and it is now considered as a primary indicator to determine their quality of life. Activity recognition is regarded as a challenging task due to the fact that each activity has their unique characteristics. It can be associated with many applications. The applications include falling detection, abnormality detection prediction of human behaviour. Activity recognition using mobile-based devices have shown to generate high quality results in a real-world setting. Due to ability of sensors in accelerometers the activity recognition shows a good potential which consumes low power and enables continuous sensing over a day. These sensors are usually embedded in various types of smartphones or it can also be strapped to human body using strip or belts. The computing capabilities of smartphones have increased in recent years and allows the function to be extended to other applications such as capturing human body movement rather than supporting only voice communications. The accelerometer sensor data in a mobile phone device helps to analyse and interrupt that user is performing activities with loads of new technologies activity recognition use to monitor human daily activities and identify anything that is unusual changes in daily life of humans .accelerometer sensor easily and publicly available but the use of a single sensor is not sufficient so more number of sensors need to accommodate may be an abstraction in movement as well as it is not too much of sensors caused really high. Due this problem researches are focusing to apply activity recognition approaches using only one accelerometer sensor to collect the body movement signal. There are also problem associated with the activity recognition approaches they include data processing features extraction methods and high performance classification support techniques .if the features are not properly selected it may degrade the activity recognition accuracy and decreases computational efficiency. Some studies applied feature extraction methods to select most significant features which classifies human activities. the dimensionality reduction process also applies to reduce the dimensionality of raw data and transform original feature to lower dimension high accuracy with in short time and real time data are several requirements of such process .This paper presents the approach to recognize activities of day to day life using publicly available accelerometer sensor data set. The issues such as signal processing feature selection dimensionality reduction and classification are also highlighted here. Some comparison process perform using machine learning classifiers which are decision tree[DT], Support vector machine [SVM], and Multi layer perceptron neural network[MLP-NN].

2. RELATED WORK

Various activity recognition approaches are proposed in different research , studies. Most of these studies composed of two categories. First categories involves the use of probabilistic model to infer the types of activities.

These method uses statistical modelling to represent the activity model structure and train the data based on large scale dataset. Hidden Markov Models (HMM), Dynamic Bayesian Network (DBN), Hierarchical Clustering and Partially Observable Markov Decision Processes (POMDP). For example in a study the activity recognition is performed using Hidden Markov Model , the features are extracted from video , camera , sensors , were genetic algorithm and Best First Search are used to select the best features that may describe the activities to be recognized. Its results shows that HMM is able to classify human activities. With feature extraction from GA with accuracy 75%. This gives an efficient way of representing activity models with probabilistic values. There are also disadvantages of these tools is that it is difficult to adapt the models in different environment due to incompleteness of training models.

The next category use classification techniques that maps input of sensory data to decide output. It use machine learning approach and techniques to extract the patterns of activities from the observed data. Random Forest (RF), Support Vector Machine (SVM) , Decision Trees (DT) and Artificial Neural Network (ANN) these are examples of classification techniques. All these methods perform a comparison process which extract the decide data with maximum accuracy and precision. The advantage of these method is handling noisy uncertain and incomplete data set. Decision Tree is used to compare the performance of these models and found that it has highest accuracy. The use of threshold based classifier is also another approach for activity recognition. It uses predefined threshold values. The activities which are generally considered are standing, sitting, walking, jumping, lying and other activities are also focused. The limit of the approach is that user needs to find the suitable threshold values as it is very sensitive to the defined activities. The fuzzy logic reasoning is also used to identify human activities. If the activities are defined in membership functions this method is sufficient. Also this only works when the fuzzy rules are clearly defined.

This paper mainly focused to analyze the activities of humans daily living based on classification techniques. Several issues are addressed in classifying the model such as feature extraction and dimensionality reduction. The comparison process is mainly highlighted so that it is performed on machine learning classifiers to find which type of features classify human activities effectively..

III. PROPOSED SYSTEM

METHODOLOGY

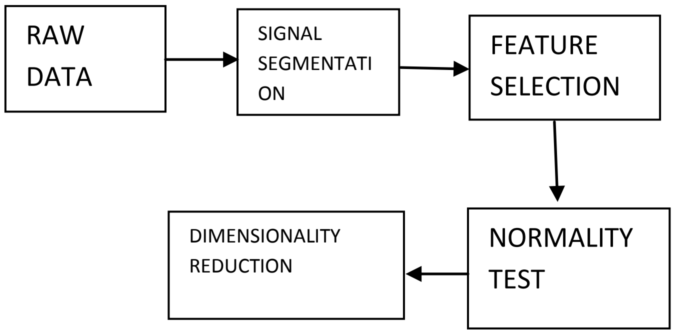
Methodology involves performing activity recognition from accelerometer sensor data embedded in a smart phone. This is divided into two subpart each part describes how data is collected and the different steps in data preprocessing method.

A. Data Collection

Publically available dataset is used here. Dataset composed of sensory data from tri axial accelerometer and a gyroscope. Because of the advantage of using multiple sensors in identifying human activities here we use only a single sensor. The accelerometer which is available in our smart phones is sufficient to determine the activity of a human. The given dataset is composed of 9 types of ADL and 4 types of false. Each signal is stored in time , acceleration values in x , y and z axis as well as labels of activity. The mean sampling rate for the signal is 87 Hz and the range of acceleration value is between 20 to -20. Here 6 activities are taken standing (STD), sitting (SIT), lying (LYI), stairsup (STU), stairs down (STN) and walking (WAL). These activities are selected as they represent the common body movements in day to day life of human. For example when we observe the activities of jumping and walking we can see different colours of graphical data is varying in x, y and z axis. This lines represents the acceleration values. After this preprocessing is applied to raw signal to generate several features.

B. Data Preprocessing

Data processing is one among the most important steps in classification technique. It consist of signal segmentation, feature selection, feature extraction and dimensionality reduction process. The following figure represents the steps in data preprocessing.



The raw data is inputed primarily then signal segmentation is performed. It is used to divide sensor signals into small time window segments in which the feature can be easily extracted in each segment. Then happens the feature selection process which allows several signal characteristics from raw sensor in data to be extracted. Here time domain and frequency domain features are widely used for the calculation of each feature. Examples of features in time domain include min, max, mean average, standard deviation, signal magnitude area(SMA) and signal vector magnitude (SVM).it's max calculated using dry axial accelerometer which consists of three axis x,y,z respectively. SVM calculated where xi is the i th sample of x axises, yi is the i th sample of y axises, zi is the ith sample of z axises accelerometer signals.

In total 60 dimensional feature vector is generated at each time.

Types of Features	Methods
Time domain	Min,max,standard deviation,signal magnitude area,signal vector magnitude,tilt angle

Freequency domain	Power spectral density,signal entropy,spectral energy

The next step which is normality test is to determined whether the accured features can be fitted in normal distribution pattern. Here determines whether it use parametric or non-parametric classification tools. There are three common test that are usually performed; Shapiro-wilk(SW), Kolmogorov-Smimov(KS) and Aandrson Drali AD(AD).This test are used to calculate the probability values(p values), and significance level for this test is 0.05. If pvalues are less than 0.05 null hypothesis is rejected.

Null hypothesis, H0; data is in normal distribution

Alternate hypothesis, H1; data is not normal.

It is concerned as one of the popular approches which reduce the dimensionality of data by converting original features into uncorrelated features. The new features are called principal compents where this are arrange accordingly to the variances and components which contribute to lowest variances are omitted. The different steps of PCA are follows;

1. Normalization of data by subtracting value

2. Co-variance matrix calculations

3. Aigen vectors and Eigenvalues calculation

4. Formation of feature vector by choosing the components.

5. New data set is derived

IV. APPLICATIONS

1. HOSPITALS

The major challenge in developing ambient intelligence environment is developing context recognition. Some systems easily recogonize contextual variables such as location, but activity recognition is much more complex. Most of the complicated cases like physical deformalities, abnormality cases, look aftering of elderly are the complex problems facing in hospitals now a days so activity recognition tries to improve the recognizing capability of human activities. Some of the studies describe an approach activity Working environment. It hidden Marko modle mostly used to classify the human behavior of data. The results indicated that hidden Markov model can correctly estimated user activity 92.6% of time and cam performs neural networks and even human observes near with work practices.

2. HOME SECURITY

We can detect possible threats and take appropriate action by monitoring activity based anomalites. Smart homes have long held the promise making our environments secure which become more productive. Individual spend the majority of their time in their home and workplace and filled that this place are their sanctuaries so in order to preserve that feeling smart homes can make use of technologies such as embedded sensor and machine learning techniques to detect identify and respond to threats. While standalone the security systems having used in home for many years it can't make the use of rich information which is available from sensors integrated throughout the home.

3. ANALYIS OF PERSON

The object detection is done by using background subtraction to detect moving object then object tracking and object classification are constructed so that different person can be differentiated which use of feature detection. Simple activity can be detected the geometric attributes to track object. These are centric and aspect ratio of identified track and manipulated .Also personalized Analaysis can be done by monitoring all activities.

4. SURVALLIANCE

It is monitoring of behavior, activities, information for the purpose of influencing managing or directing .This is made easy by the activity recognition using accelerometer sensors embedded in smartphone with the machine learning classifiers.

5. ANTI CRIME SECURITY

The data collected can be used by detect gunfire and pin point where the gunshots came from. This helps to detect the hidden ones which threaten life of people.

V. RESULT

The result of normality test are presented here

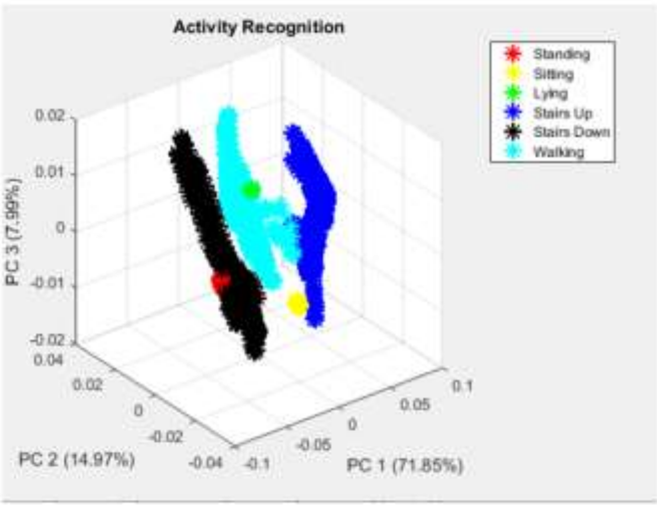
Types of features	AD	SW	KS
Mean	P-value<0.05	P-value<0.05	P-value<0.05
Standard deviation	P-value<0.05	P-value<0.05	P-value<0.05
SMA	P-value<0.05	P-value<0.05	P-value<0.05
SVM	P-value<0.05	P-value<0.05	P-value<0.05
PSD	P-value<0.05	P-value<0.05	P-value<0.05

From the table it can be seen that some features are taken from the time and frequency domain have probability values less than 0.05. As a result the null hypothesis can be rejected and it can be concluded that acquired features are not fitted in normal distribution with full confidence. So here non parametric tools can be choosen for the classification process. The result of dimensionality reduction process using PCA will be a three dimensional plot which is used to calculate the most important principle component values. This all contain useful informations and clearly shows each activity is successfully clustered into their own groups. Next the result of comparison process is obtained. It is calculated in terms of precision, recall f - score and accuracy. The original features are extracted from raw data and PCA based features are also calculated. Time domain and frequency domain features are also compared.

The results obtained from the comparison of classification performance between original features and feature after dimensionality reduction shows that out of 3600 samples. Each activity contain 600 samples of data. It can be observed that calculated rate are higher than 90%. MLPNN gives the best results compared to all other classifiers.

On the other hand it can be seened as SVM gives relatively worst results with an accuracy of 91%

Here we can see that the total average accuracy is increased by 4% . When dimensionality reduction using PCA is applied to the original features. So we can say that PCA has reduced the data dimensionality and only the highest variances is explained (PC1 , PC2 & PC3) are chosen as the data set for classification process.



The comparison of classifiers between features are divided into time domain and frequency domain. Frequency domain features shows higher total average accuracy compared to time domain features. Based on the result it can be concluded that frequency domain features give more meaningful data representation when it is compared to time domain features. The reason for this is frequency domain uses discrete forrier transform to transform the features in the time domain. This gives better characteristics of the signal and help to categorize well different human activities. The machine learning classifiers such as MLP-NN shows the example of confusion matrix. Its input is based on data after dimensionality reduction process. It has 6 classes of activities standing, sitting , lying , stairs up , stairs down , walking. From that table it can be seen that the activities are accurately recognized with 100% accuracy.

Confusion matrix

	STD	SIT	LYI	STU	STN	WAL
STD	186	0	0	0	0	0
SIT	0	178	0	0	0	0
LYI	0	0	166	0	0	0
STU	0	0	0	167	0	0
STN	0	0	0	0	194	0
WAL	0	0	0	0	0	189

3. CONCLUSIONS

The paper presents the apparoch of recognizing the activities of daily living based on commonly available accelerometer sensor dataset. This data set is usually uses an accelerometer sensor which is embedded in smartphone. A number of feature from time domain and frequency domain are extracted from the raw accelerometer signal. Principal component Analaysis is performed on the original features which distinguish low and high variances .This due to the dimensionality of data. This approaches evaluated by comparing the precision, recall, F-score and accuracy of four different types of machine learning classifiers. The normality test results proven that data is not in normal distribution therefore non parametric classification tools are used the classify the activities. PCA based features gives the constrable improvement which improve the recognition rate rather than using original features. Features that are selected from frequency domain shows higher accuracy rather than time domain features. The use of multiple sensors is not practical and make as many difficulties. The problem of movement is major difficulty and is not applicable for long term wearing. So the paper uses only one accelerometer sensor which performes significantly better. Many types of security cameras application are used now a days to detect the activity recognition but it is not possible whole times. If regular monitoring is not happened, this purposes may be failled. So activity recognition using accelerometer sensor embedded in a smartphone which is commonly available is uesd which is very much effective recognizing human activities and the data and results will be precise the correct when machine learning classifiers applied.

REFERENCES

[1] S. Chernbumroong, A. S. Atkins, and H. Yu, “Perception of Smart Home Technologies to Assist Elderly People,” 4th Int. Conf. Software, Knowledge, Inf. Manag. Appl. (SKIMA 2010), no. March 2016, pp. 1–7, 2010.

[2] S. Iwarsson, V. Horstman, and U. Sonn, “Assessment of Dependence in Daily Activities Combined with a Self-rating of Difficulty,” Journal of Rehabiitation Medical, vol. 41, no. 3, pp. 150–156, 2009

[3] T. Shi, X. Sun, Z. Xia, L. Chen, and J. Liu, "Fall Detection Algorithm Based on Triaxial Accelerometer and Magnetometer," no.2, May, 2016

. [4] F. Abdali-Mohammadi, M. Rashidpour, and A. Fathi, "Fall Detection using Adaptive Neuro-fuzzy Inference System," International Journal of Multimedia Ubiquitous Eng., vol. 11, no. 4, pp. 91–106, 2016.

[5] E. Sriperumbudur, "An Unsupervised Pattern Clustering Approach for Identifying Abnormal User Behaviors in Smart Homes," Int. J. Comput. Sci. Netw., vol. 2, no. 3, pp. 115–122, 2013.

[6] M. Novák and F. Jakab, "Approaches to a Problem of Anomaly Detection in User Behaviour in a Smart Home [17] R. Cilla, M. A. Patricio, J. García, A. Berlanga, and J. M. Molina, "Recognizing Human Activities from Sensors using Hidden Markov Models Constructed by Feature Selection Techniques," Algorithms, vol. 2, no. 1, pp. 282–300, 2009.

[18] K. G. Manosha Chathuramali and R. Rodrigo, "Faster Human Activity Recognition with SVM," in International Conference on Advances in ICT for Emerging Regions, ICTer 2012 - Conference Proceedings, 2012, pp. 197–203. [19] M. N. S. Zainudin, N. Sulaiman, N. Mustapha, and T. Perumal, "Activity Recognition based on Accelerometer Sensor using Combinational Classifiers," pp. 68–73, 2015.

[20] G. Vavoulas, M. Pediaditis, E. G. Spanakis, and M. Tsiknakis, "The MobiFall dataset: An Initial Evaluation of Fall Detection Algorithms using Smartphones," 13th IEEE Int. Conf. Bioinforma. Bioeng., no. November, pp. 1–4, 2013.

[21] L. Chen, C. Nugent, and G. Okeyo, "An Ontology

based Hybrid Approach to Activity Modeling for Smart Homes," IEEE Trans. Human-Machine Syst., vol. 44, no. 1, pp. 92–105, 2014.

[22] Z. Zeng and Q. Ji, "Knowledge Based Activity Recognition with Dynamic Bayesian Network," Network, pp. 532–546, 2010

Environment," in Proceedings of the Faculty of Electrical Enginering and Informatics of the Technical University of Kosice, 2011, pp. 1–5.

[7] S. Bourobou and Y. Yoo, "User Activity Recognition in Smart Homes Using Pattern Clustering Applied to Temporal ANN Algorithm," Sensors, vol. 15, no. 5, pp. 11953–11971, 2015.

[8] S. Mahmoud, A. Lotfi, and C. Langensiepen, "Behavioural Pattern Identification and Prediction in Intelligent Environments," Appl. Soft Comput. J., vol. 13, no. 4, pp. 1813–1822, 2013.

[9] C. L. R. Mcghan, A. Nasir, E. M. Atkins, A. Aerospace, and A. Arbor, "Human Intent Prediction Using Markov Decision Processes," Infotech AIAA Calif.,no.June,pp.

[10] I. Fatima, M. Fahim, Y. K. Lee, and S. Lee, "A Unified Framework for Activity Recognition-based Behavior Analysis and Action Prediction in Smart Homes," Sensors (Switzerland), vol. 13, no. 2, pp. 2682–2699, 2013.

[11] T. Sztyler, "On-body Localization of Wearable Devices : An Investigation of Position-Aware Activity Recognition," 2016.

[12] F. Attal, S. Mohammed, M. Dedabrishvili, F. Chamroukhi, L. Oukhellou, and Y. Amirat, "Physical Human Activity Recognition Using Wearable Sensors," Sensors, vol. 15, no. 12, pp. 31314–31338, 2015.

[13] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. P. Cardoso, "Preprocessing Techniques for Context Recognition from Accelerometer Data," Pervasive Ubiquitous Computing., vol. 14, no. 7, pp. 645–662, 2010

. [14] J. Fu, C. Liu, Y. Hsu, and L. Fu, "Recognizing Context - aware Activities of Daily Living using RGBD Sensor," Iros2013, pp. 2222–2227, 2013.

[15] W. Xiao and Y. Lu, "Daily Human Physical Activity Recognition Based on Kernel Discriminant Analysis and Extreme Learning Machine," Math. Probl. Eng., vol. 2015, 2015.

[16] I. Farkas and E. Doran, "Activity Recognition from Acceleration Data Collected with a Tri-axial Accelerometer," Acta Tech. Napocensis - Electron. Telecommun., vol. 52, no. 2, pp. 38–43, 2011.

[17] R. Cilla, M. A. Patricio, J. García, A. Berlanga, and J. M. Molina, "Recognizing Human Activities from Sensors using Hidden Markov Models Constructed by Feature Selection Techniques," Algorithms, vol. 2, no. 1, pp. 282–300, 2009.

[18] K. G. Manosha Chathuramali and R. Rodrigo, "Faster Human Activity Recognition with SVM," in International Conference on Advances in ICT for Emerging Regions, ICTer 2012 - Conference Proceedings, 2012, pp. 197–203.

[19] M. N. S. Zainudin, N. Sulaiman, N. Mustapha, and T. Perumal, "Activity Recognition based on Accelerometer Sensor using Combinational Classifiers," pp. 68–73, 2015.

[20] G. Vavoulas, M. Pediaditis, E. G. Spanakis, and M. Tsiknakis, "The MobiFall dataset: An Initial Evaluation of Fall Detection Algorithms using Smartphones," 13th IEEE Int. Conf. Bioinforma. Bioeng., no. November, pp. 1–4, 2013.

[21] L. Chen, C. Nugent, and G. Okeyo, "An Ontology based Hybrid Approach to Activity Modeling for Smart Homes," IEEE Trans. Human-Machine Syst., vol. 44, no. 1, pp. 92–105, 2014.

[22] Z. Zeng and Q. Ji, "Knowledge Based Activity Recognition with Dynamic Bayesian Network," Network, pp. 532–546, 2010