E-CARE: ANDROID APPLICATION FOR HEALTH MONITORING AND HUMAN PHYSICAL ACTIVITY CLASSIFICATION

[1]kadambari Hadke, [2]Prerna Abhang, [3]Prachi Golhar, [4]Sagarika Patil
[1]Student, [2]Student, [3]Student, [4]Student
[1]Computer Science Engineering,
[1]Pune Institute Of Computer Technology, Pune, India

Abstract*—In this fast moving world obesity is one of the serious problems for the health of worldwide population. Social Interactions on mobile phone via Internet are major cause of lack of physical activities. Mobile Sensors can be used to collect health related information from the person. Healthy Life Style plays vital role in improving resistivity power of the person. Using Accelerometer Sensor Daily Activities such as Walking, Standing, Jogging, Sitting, Walking Upstairs, Walking Downstairs are classified to maintain Healthy Life Cycle. KNN Classification Algorithm will be used to categorize these activities. Later on these activities would help in making a diet plan for user using decision tree algorithm. If any problem then user may contact doctor. The diet plan would be approved by doctor.

***Index Terms*—Smartphone, Acceleration, Physical Activity Classification, Healthcare, Diet Plan**

I. INTRODUCTION

People are generally considered obese when their body mass index (BMI), a measurement obtained by persons weight by the square of person's height is over 30kg/m^2 .Worldwide obesity is nearly doubled after 1980. World Health Organization (WHO) says that more than 10% of population is obese. Obesity is also related to large number of chronic diseases [1]. Recently, it is found that the number of years lived with obesity is directly proportional to the risk of mortality. According to world health organization, about 2.8 million people are dying every year due to obesity related diseases. To reduce the problem of obesity, preventative efforts include proper diet and enhanced daily physical activities. Lot of research has been done to optimize the diet and exercise plan to reduce the obesity in adults and children. It is reported in the literature that both diet and physical activity are important factors [3–5]. Many nutritionist and doctors monitor the physical activity of patients by self-filled questionnaires to assess the amount of physical activity. Physical activity index based on the questionnaires are also proposed to assess different level of activeness of the people.

In this Project we are using in built Mobile Sensor to maintain healthy lifestyle for people efficiently. We will try to overcome above existing Systems problems. In this project we are trying to eliminate the need of wearable. Classification of activities such as Sitting, Jogging, Walking, Walking Upstairs, Walking Downstairs, Running, and Standing done based on readings. Emergency alarm is being sent. On the Basis of activities we are trying to calculate Calories how much they have burn. And then proper diet will be suggested to the Patient.

II. MATERIALS AND METHODS

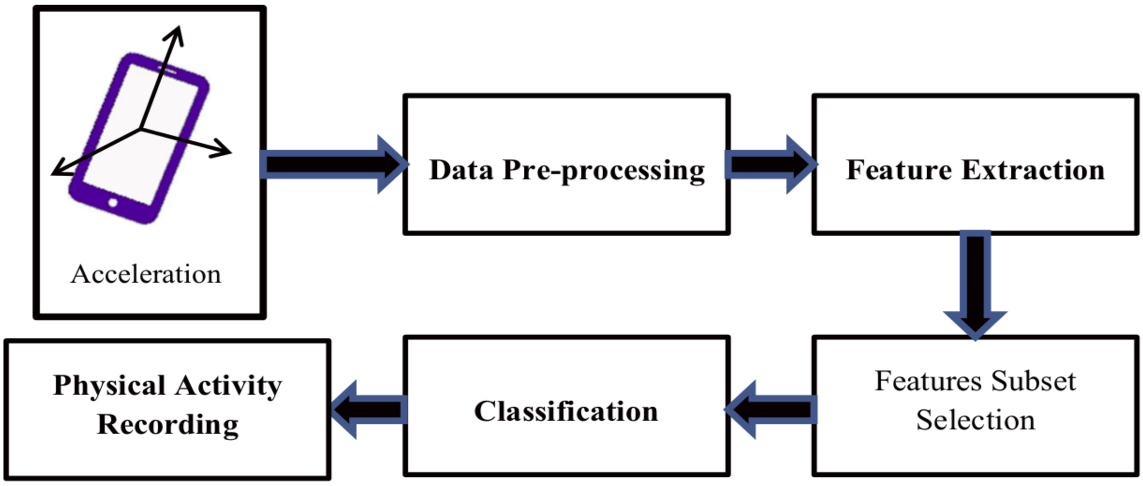

Fig. Physical activity recognition using Smartphone acceleration sensor

A. Data Collection :

In this paper, the data we used was collected from accelerometer sensors present inside the mobile phones. The user is not required to tie any wearable device with him; instead he can freely perform any activities with the mobile phone in his pocket. In our experiment, user performed six different physical activities like walking, upstairs-downstairs, standing, sitting, jogging etc. As shown in fig.1, Accelerometer sensor returns data in x, y, z-direction of all six physical activities. Activity of jogging produces periodic movements in x-direction and it has high amplitude as compared to the walking activity. The variation of signal is very little in case of sitting and standing activity. Walking upstairs and downstairs are somewhat similar to walking with lesser periodicity.

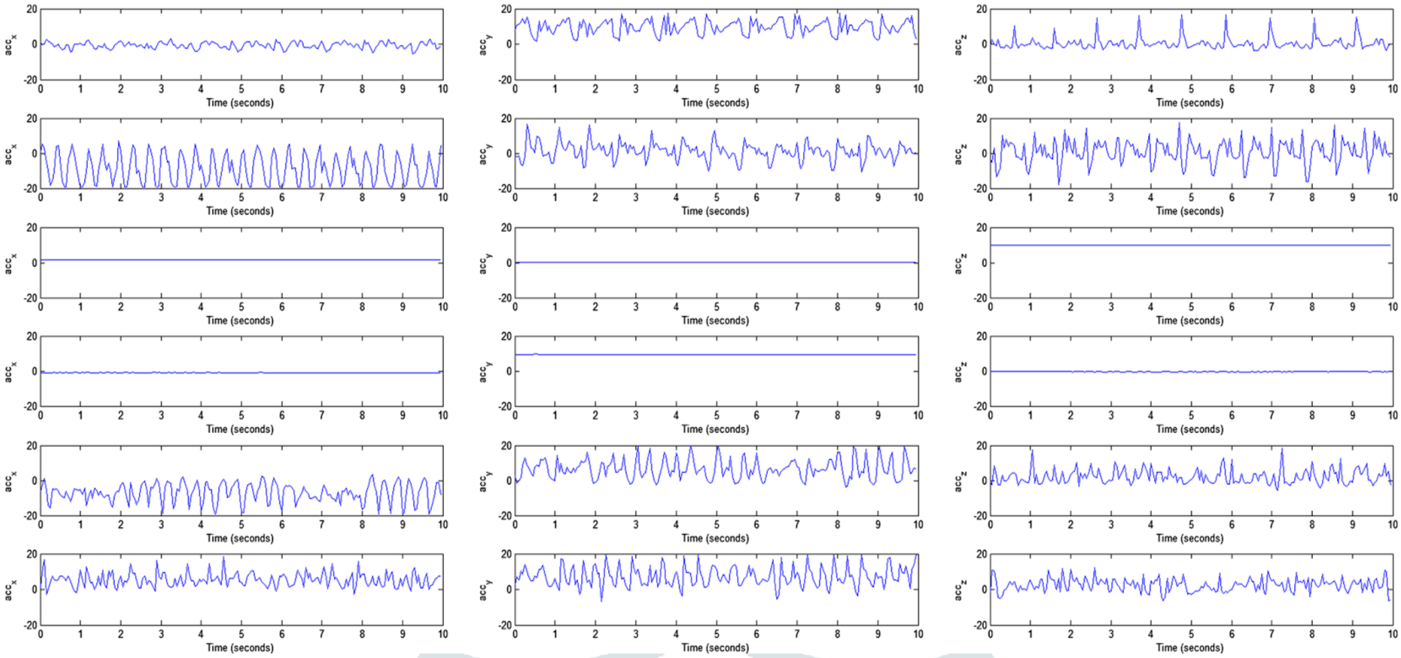

Fig.1 the activity signals collected by the accelerometer sensor

B. Acceleration data preprocessing:

Before we move on to calculating any feature, the data we get from accelerometer needs to be preprocessed in order to reduce noise using median filter or order n in each dimension separately. A window of wt seconds (fs × wt samples) is used to calculate the feature set for a particular activity. Here, fs is the sampling frequency of the acceleration data.

C. Feature Extraction:

Every window of wt second consists of acceleration in ax, ay and az dimension. Acceleration given by accelerometer sensor in each direction is the sum of gravitational, 'g' and body, 'b' accelerations. Thus, a low pass filter is used to separate the acceleration signal into gravity (agx, agy and agz) and body acceleration(abx , aby and abz).

D. Feature Subset Selection:

It is important that we analyze all the features that we have extracted and select only those minimum numbers of features which contribute more in the correct classification of the physical activities. This Feature subset selection helps to reduce the processing cost and it also improves the performance of the model.

E. Classification of physical activity:

The K-Nearest Neighbor classification algorithm is used to classify the activities. The KNN classification algorithm is widely used classification algorithm. The data is decided based the class labels of the neighboring instances. This Algorithm gives better accuracy as compared to other classification algorithms.

III. EXPERIMENTAL ENVIRONMENT

As discussed above the acceleration data being stored in x ,y , z directions are divided into instances by sliding window $_t$ for some time. Features described in above sections are calculated for some specific time and feature sets of instances are created where these instances contain 105 features**.**

Table1 Classification results by KNN onFS1

Activity	T-P rate	F-P rate	Precision	Recall	F-measure
Walking	0.998	0.004	0.994	0.998	0.996
Jogging	0.998	0.001	0.998	0.998	0.998
Sitting	0.998	0	1	0.989	0.995
Standing	1	0.001	0.988	1	0.994
Upstairs	0.983	0.002	0.988	0.983	0.985
Downstairs	0.976	0.001	0.99	0.976	0.983
Average	0.994	0.002	0.994	0.994	0.993

The results obtained after feature selection is classification being completed done by Weka software. Classification of feature set is done by using KNN algorithm where value of K is 3. Results given in Table1 are calculated by 10 cross fold validation. 99.4% is overall classification accuracy. TP is true positive rate and FP is false positive rate.

The proportion of instances that belong to a class by total instance classified by classifier is the proportion whereas proportion of instances classified by the classifier as belong to this class is nothing but recall. Combination of recall and proportion is f-measure and mathematically written as,

$F_{measure}$=2*Precession*Recall/Precesion+Recall

Table2: Confusion matrix of KNN

	Walking	Jogging	Sitting	Standing	Upstairs	Downstairs
Walking	8347	0	0	0	9	4
Jogging	2	6718	0	0	5	0
Sitting	0	0	1104	11	1	0
Standing	0	0	0	885	0	0
Upstairs	18	11	0	0	2306	10
Downstairs	31	1	0	0	14	1849

0.993 (99.3% in percentage) as F-measure is good performance of feature set for all activities. Table 2 shows confusion matrix for KNN classifier on FS1 where instances of walking and jogging are confused with instances of upstairs and downstairs and vice-versa.

Table 3 Reduction of feature subset by CFS on FS1

	Reduced Scatter Search FS2(30)			Scatter Search FS3(41)			Subset Size FS FS4(42)		
Activity	***Precision***	***Recall***	***F-measure***	***Precession***	***Recall***	***F-measure***	***Precession***	***Recall***	***F-measure***
Walking	0.972	0.993	0.982	0.980	0.996	0.988	0.980	0.995	0.987
Jogging	0.997	0.995	0.996	0.997	0.997	0.997	0.997	0.997	0.997
Sitting	0.998	0.995	0.996	0.998	0.994	0.996	0.998	0.995	0.996
Standing	0.993	0.997	0.995	0.993	0.997	0.995	0.994	0.997	0.995
Upstairs	0.978	0.953	0.965	0.979	0.959	0.969	0.980	0.960	0.970
Downstairs	0.963	0.908	0.935	0.976	0.933	0.954	0.969	0.930	0.949
Average	0.982	0.982	0.982	0.986	0.986	0.986	0.986	0.986	0.986

There may be some attributes redundant in feature set FS1. Therefore correlation based feature selection (CFS) is used to remove the redundant and irrelevant features and reduce feature set. Some search methods are scatter search (SS), reduced scatter search (RSS) and subset sized forward selection (SSFS) used to search the best feature subset. Table 3 shows classification results of KNN classifier on feature subsets form SS, RSS and SSFS. Average values of precession, recall and f-measure are almost equal for all search methods. Reduced scatter search method provides with other feature subset FS2 that has less number of features. Table 4 shows confusion matrix of classification results using KNN on FS2 with 10 fold validations. Instances of various activities like walking, walking down stairs, walking upstairs is confused and there are many similarities in acceleration of these three activities. The results published [13, 14] are evident for these observation.

Table 5shows comparison of performance of three classifiers using feature subset FS2 having 30 features. J48 classifier [12] is base classifier in rotation forest classifier and principal component analysis (PCA) is the extraction method. 10 trees are constructed from 5 random features in random forest classifier. Random forest takes less time than rotation forest to build the model. But KNN has more time and space complexity for searching the class of a query data point than random forest and rotation forest. But overall classification accuracy of KNN for 10 cross folds is better than rotation forest and random forest (Table5).

In query classification whole dataset is used as representative instances in KNN based classification. Therefore it is important to remove redundancy or less significant instances from training dataset to reduce size. Detrimental Reduction Optimization Procedure 2(DROP2) is one of the algorithms to select significant instances with respect to classification proposed by Wilson and Marteinz. For set of instances S algorithm starts by taking all the instances from the set S into T and then removes an instance P from the set T if at least as man y of its associates in the original set T including the instances already removed from Tare classified correctly without P. This procedure done for all instances in the set T. The division of FS2 is training dataset and testing dataset. 80% of dataset included in training dataset and 20% of data set instances are included in testing set. On the training dataset and instances DROP2 pruning algorithm is used and stored FS5 dataset. Table 7 summarizes classification results.

Table 7 shows confusion matrix of KNN classifier applied on the training dataset using the pruned dataset FS5. The instances of all six classes are classified correctly but with confusion among upstairs, downstairs and walking. Comparison of similar experimental setups published in different papers. About 80% accuracy is shown by Maurer et al[10] where a sensor placed in the trouser pocket for six types of activities . SVM classifier was used on 66 feature sand achieved F-measure equals to 93%. Overall classification accuracy was achieved as 90.8%

Acceleration sensor was placed on the waist belt and classified eight activities (Sitting, Standing, Lying, Walking, Sit-to-Stand, Stand-to-Sit, Stand-to-Lie and Lie-to-Stand) this experiment was performed by Allen et al[11] and achieved accuracy of 91%. Small pilot study was conducted by Mi et al [58] on five activities (Sitting, Standing ,Lying ,Walking, Running) to get classification accuracy of 99%.

Our framework produced much better than results as compared to published results. 99.4% overall accuracy was achieved which is highest. Pruning is used to eliminate redundancy and speed up the time complexity. The features calculated are in time domain. So the time complexity of KNN classifier with feature subset of only 30 features and 1689 representative prototypes will be space and time efficient.

Table4: Confusion matrix of KNN classifier on FS2

	Walking	Jogging	Sitting	Standing	Upstairs	Downstairs
Walking	8300	2	0	0	18	40
Jogging	20	6698	0	0	9	3
Sitting	0	0	1110	6	0	0
Standing	0	0	2	882	2	0
Upstairs	72	15	0	0	2235	23
Downstairs	148	4	0	0	22	1721

Table 5: Classification results of three classifiers on FS2

	KNN			Random Forest			Rotation Forest		
Activity	*Precision*	*Recall*	*F-measure*	*Precession*	*Recall*	*F-measure*	*Precession*	*Recall*	*F-measure*
Walking	0.972	0.993	0.982	0.923	0.988	0.954	0.899	0.990	0.942
Jogging	0.997	0.995	0.996	0.987	0.992	0.989	0.984	0.988	0.986
Sitting	0.998	0.995	0.996	1.000	0.990	0.995	0.997	0.993	0.995
Standing	0.993	0.997	0.995	0.991	1.000	0.996	0.992	0.998	0.995
Upstairs	0.978	0.953	0.965	0.937	0.856	0.895	0.918	0.809	0.860
Downstairs	0.963	0.908	0.935	0.931	0.724	0.815	0.921	0.629	0.748
Average	0.982	0.982	0.982	0.952	0.952	0.950	0.939	0.938	0.935

Table 6: **Classification results of KNN (DROP2 based reduction on FS5)**

		DROP2 Pruning			DROP2 Pruning		
Activity	*%age selected*	*Precision*	*Recall*	*F-measure*	*Precession*	*Recall*	*F-measure*
Walking	10.31%	0.978	0.973	0.975	0.963	0.959	0.954
Jogging	5.49%	0.993	0.995	0.996	0.986	0.985	0.989
Sitting	4.95%	0.984	0.995	0.996	0.996	0.990	0.995
Standing	5.87%	0.991	0.98	0.986	0.989	0.995	0.992
Upstairs	24.18%	0.932	0.946	0.939	0.888	0.882	0.885
Downstairs	29.44%	0.903	0.913	0.908	0816	0.841	0.828
Average	11.60%	0.972	0.971	0.971	0.953	0.953	0.953

Table 7: **Confusion matrix of KNN classifier on the training set using FS**

	Walking	Jogging	Sitting	Standing	Upstairs	Downstairs
Walking	8300	2	0	0	18	40
Jogging	20	6698	0	0	9	3
Sitting	0	0	1110	6	0	0
Standing	0	0	2	882	2	0
Upstairs	72	15	0	0	2235	23
Downstairs	148	4	0	0	22	1721

IV. CONCLUSION

Physical activity monitoring is very important in day-to-day life. In this paper, classification results are presented for six types of physical activities Currently, The system using accelerometer sensor of mobile to classify the human physical activities such as walking, upstairs, downstairs, jogging, seating and standing to maintain healthy lifecycle. Using KNN algorithm for classification of activities and Decision algorithm i.e. ID3 is used for training data for suggesting diet plan**.** A proper diet plan would be created by using information at registration and these activities being classified.

It was found that KNN classifier is the best classifier than other classifiers in terms of accuracy. It is shown that this classifier can produce more than 95 % classification accuracy.

REFERENCES

[1] N. Nirwal, Hopeful hearts a mobile healthcare application. IEEE, 2014.

[2] R. S. Pressman, Software Engineering (3rd Ed.): A Practitioner's Approach. New York, NY, USA: McGraw-Hill, Inc., 1992.J.ClerkMaxwell,ATreatiseonElectricityandMagnetism,3rded.,vol.2.Oxford:Clarendon,1892,pp.68–73.

[3] S. Tyagi, A conceptual framework for IOT based Healthcare System using Cloud computing. IEEE 6th International conference, 2016.K.Elissa, "Titleofpaperifknown," unpublished.

[4] P. Li, An automatic user adapted physical activity classification methods using Smartphone. IEEE, 2016.

[5] S. Zhang, Effective ways to use IOT in the field of medical and smart healthcare. IEEE, 2016.

[6] Better Physical Activity Classification Using Smartphone Acceleration Sensor, September 2014 DOI: 10.1007/s10916-014- 0095-0 · Source: PubMed CITATIONS

[7] Jerk-based feature extraction for robust activity recognition from acceleration data.Wilhelmiina H¨am¨al¨ainen, Mikko J¨arvinen, Paula Martiskainen, Jaakko Mononen .Department of Biosciences University of Eastern Finland Kuopio, Finland whamalai@uef.fi

[8] Mobile Phone Sensors in Health Applications.Evgeny Stankevich, Ilya Paramonov, Ivan Timofeev P. G. Demidov Yaroslavl State University Yaroslavl, Russia fstankevich.evg, skat.setg@gmail.com, ilya.paramonov@fruct.org

[9] Design and implementation of Android-based health and healthcare system. Huifang DENG and Sihan CHEN

[10] Maurer, U., et al. Activity recognition and monitoring using multiple sensors on different body positions. in Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on. 2006. IEEE

[11] Allen, F. R., et al., Classification of a known sequence of motions and postures from accelerometry data using adapted Gaussian mixture models. Physiological Measurement 27(10):935, 2006.

[12] Quinlan, J.R., C4. 5: programs for machine learning. Vol. 1. 1993: Morgan kaufmann.

[13] Kwapisz, J. R.,Weiss, G. M., and Moore, S. A., Activity recognition using cell phone accelerometers. ACM SigKDD Explorations Newsletter 12(2):74–82, 2011.

[14] Lester, J., T. Choudhury, and G. Borriello, A practical approach to recognizing physical activities, in Pervasive Computing. 2006, Springer. p. 1–16.