



Analysis Based on Estimating Heart Rate While Classifying Activities Using Wearable Sensor Data

Priyanka Kolluri¹, Manasa Pittala²

Student, Department of ECE, VNRVJIET, Hyderabad, India¹

Student, Department of ECE, VNRVJIET, Hyderabad, India²

Abstract: In the era of smart life, wearable devices are revolutionizing health care. Estimating the heart rate of a person to monitor their health and fitness is an integral part of using wearable devices that contain Photoplethysmography (PPG) sensors. Though the PPG sensor gives an easier estimation of heart rate when compared to electrocardiography (ECG), the motion artefacts can act as an impediment affecting the accuracy and thus 3D accelerometer sensor is also used to combat this. This paper proposes a novel approach in order to classify and estimate heart rate while performing activities using machine learning models. The results of which show that an accuracy of 96.67% is obtained when the Random Forest classifier is used followed by other ensemble classifiers like Bagging and AdaBoost classifiers.

Keywords: Wearable devices, PPG sensor, 3D Accelerometer, heart rate estimation, activities.

I. INTRODUCTION

Cardiovascular health is an indispensable part of human life. Any improvement in the health or fitness regime is directly linked with estimating the heart rate and monitoring it continuously. Photoplethysmography (PPG) sensors use low-intensity infrared light to estimate heart rate. When light travels through the skin, it gets absorbed more by the blood than the tissues. Thus, the PPG sensors can detect the change in the blood flow with a change in the light intensity using motion tolerant technology. This mechanism is really helpful and is being used in all the latest technological gadgets and medical devices to maintain the health and fitness of people. The electrocardiography (ECG) sensor which uses electric signals to measure heart rate is not preferred to the PPG sensor because it is inconvenient in day-to-day life while performing activities. The Photoplethysmography (PPG) sensor shows less efficiency while producing a reading due to motion artefacts so accelerometer data is also used along with it to increase its accuracy and estimation rate. This paper uses the PPG-DaLiA dataset [1] and proposes a methodology using machine learning models to predict the heart rate produced during different activities performed by various subjects in close to real-life conditions.

II. LITERATURE SURVEY

The PPG-DaLiA dataset refers to heart rate-based measurements during the continuous execution of various activities [2]. There are multiple authors who worked on this dataset proposing various methodologies in order to produce results with lesser noise, more feasibility, low cost to benefit the healthcare domain. A.P. Psatha [3] applied Fast Fourier Transform (FFT) on the dataset and obtained best model accuracy of 92.8% for Bagged Trees. H. Gao and others [4] assessed signal quality using LSTM network to separate clean and noisy signals. M. Wójcikowski and B. Pankiewicz [5] applied signal conditioning and proposed an algorithm for low-power and less computational cost based on input quality of the signal. P. H. Charlton [6] demonstrated a pilot study evaluating the ability to receive data from PPG sensors and concluded that pulse wave technology may be useful. P. Casale and others [7] introduced a chest wearable device with 319 features and 90 % accuracy was obtained to classify activities like walking, climbing stairs etc.

III. METHODOLOGY

This paper proposes a methodology using PPG-DaLiA dataset to estimate heart rate while recognizing different activities. The data is pre-processed and features are extracted before dividing the dataset into testing and training files. The testing and training data are then given to different machine learning models for classification, recognition of activities, and estimation of heart rate.

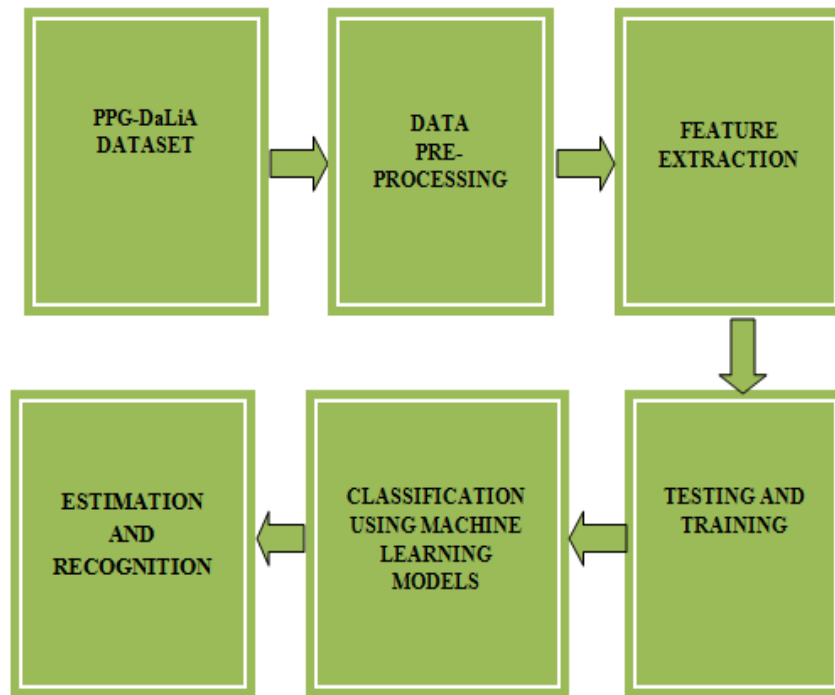


Fig. 2 Block diagram for estimation of heart rate while recognizing activities

A. PPG-DaLiA Dataset:

The experiment is based on the PPG-DaLiA dataset which contains raw data taken from a chest-worn device called RespiBan professional and a wrist-worn device called Empatica. The ECG signal data, respiration signal and 3D accelerometer sensor data are acquired from RespiBan professional whereas temperature, electro-dermal activity (EDA), 3D accelerometer sensor and photoplethysmography (PPG) sensor data are captured from the Empatica. The accelerometer sensor is embedded into the wearable device which has a PPG sensor to compensate for motion artefacts. The ECG signal data is taken as the ground truth reference to estimate the heart rate. The data is collected from 15 subjects which include 7 men and 8 women around the age groups of (30.60 ± 9.59) years. These subjects were performing 8 different activities in close to real-life conditions like sitting, ascending and descending stairs, table soccer, cycling, driving a car, lunch break, walking, and working labeled from 1 to 8 respectively. There is also an activity ID labeled 0 which corresponds to transient period between any 2 activities.

B. Data Pre-Processing:

In this block of data pre-processing, the primary parts are sampling and windowing. The signal data is segmented with a sliding windowing width of 8 seconds with a sliding windowing shift of 2 seconds. The heart rate is entirely determined in the 8 seconds window with a 2 seconds shift. This 8/2 window segmentation produces a ground truth of heart rate when extracted from an ECG signal. Data from the PPG sensor is sampled at a frequency of 64Hz and the wrist accelerometer signal data is sampled at a rate of 32 Hz. The synchronized and labeled data is then stored in different frequencies. Thus, it is necessary to process the data under one target selected frequency of 4Hz where the activity can be recorded.

C. Feature Extraction:

In this feature extraction block, an important part is to extract the R peak value in the correct time period. R-peak is obtained from an ECG wave and counting the number of R-peaks in the time of 0.25 seconds which directly determines the recognized activity at 4Hz of frequency. The data obtained from 15 subjects have different attributes like age (years), height (cm), weight (kg), skin type (taken from Fitzpatrick scale) and fitness level representing a scale value from 1 to 6, which determines the type of exercise that several subjects are engaged in, where 1 refers to the subject being active for less than a month and 6 referring to the subject being active for 5 to 7 times a week.

D. Testing and Training:

The feature dataset is then divided into training and testing data in the ratio of 75 percent to 25 percent respectively. These data are labeled according to the behavior of different participants, which makes them unique.



IV. EXPERIMENT RESULTS

The training and testing data are given as input to different machine learning models for estimating heart rate based on the classification of activities performed. The different classifiers like Random Forest, Decision tree, AdaBoost, and Bagging classifiers are used. The accuracy is determined based on the confusion matrix generated by the classifiers when estimating and predicting the heart rate while performing activities [8].

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{False Positive} + \text{True Negative} + \text{False Negative}}$$

A. Heart Rate Estimation While Recognizing Activities Using Random Forest Classifier:

Random forest classifier is one of ensemble classifiers other than Bagging and AdaBoost classifiers. It is used to solve classification problems by using multiple decision trees and takes the output which consist most number of votes. In this experimentation, the feature dataset is given as an input to the random forest classifier and the highest accuracy of 96.67% is obtained to assess heart rate and predict activity [9]. The confusion matrix is generated from the true labels and predicted labels of all activities. This can be used for evaluating accuracy and performance of the classifier. The Fig. 2 shows the confusion matrix generated when random forest classifier is used with highest accuracy. The label IDs 1-8 represent activities performed and 0 represents a transient period that exists between each activity.

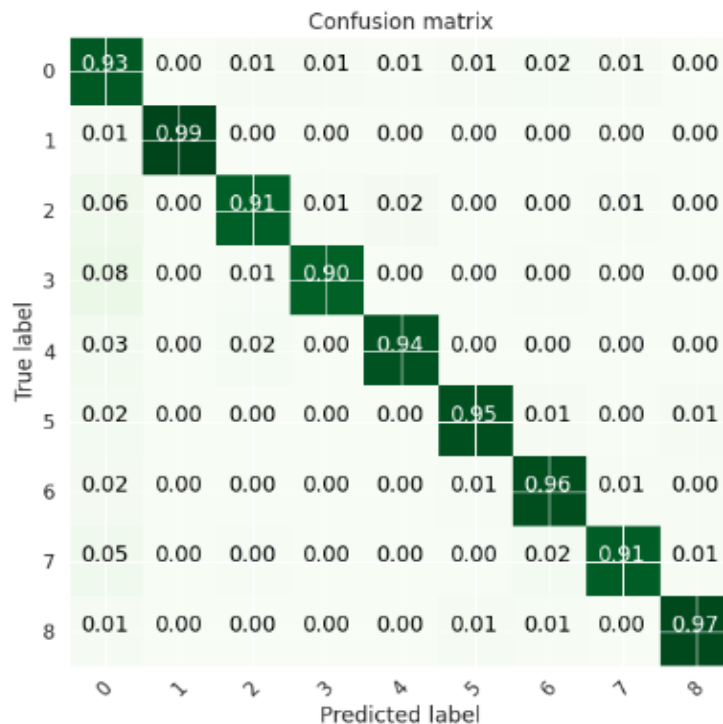


Fig. 2 Confusion matrix for classification of activities using random forest classifier

The other classifiers are also used for estimation and prediction but they produce lesser accuracies when compared to Random Forest classifier.

V. ANALYSIS BASED ON ESTIMATION AND RECOGNITION

The classifiers generated various output accuracies when predicting heart rate based on activities. The highest accuracy of 96.67% is obtained when random forest classifier is applied on the feature dataset followed by the ensemble classifiers like bagging classifier with 96.35% accuracy and AdaBoost classifier with 94.43% respectively [10]. The least accuracy of 94.4% is obtained when a Decision tree classifier is used. The Fig. 3 shows the graphical representation based on the comparison among various classifiers used for estimation of heart rate while performing different activities using PPG sensor and 3D Accelerometer sensor data.

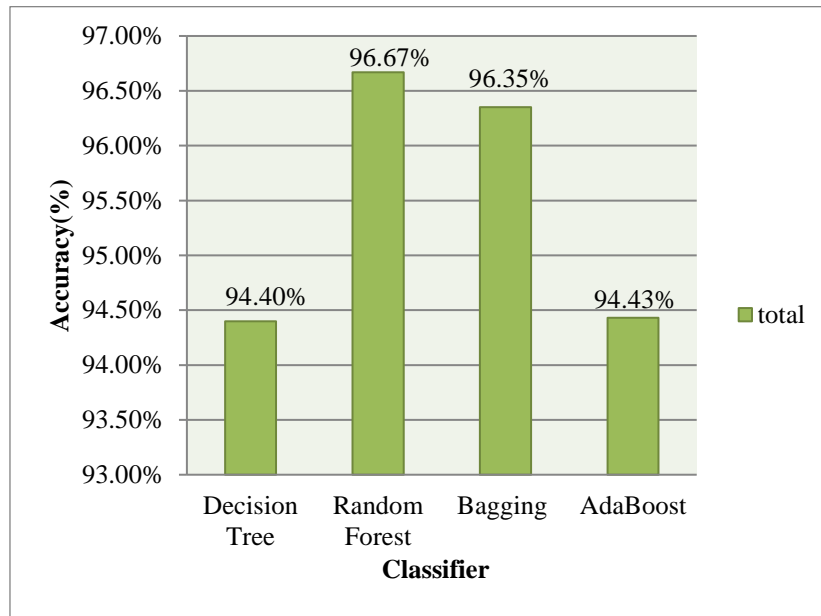


Fig. 3 Analysis based on recognition of activities using different classifiers

VI. CONCLUSION AND FUTURE SCOPE

This paper proposes a methodology for assessing of a heart rate while performing activities using PPG sensors which is really promising. The PPG sensors and 3D accelerometer sensors are practically cost effective and also more convenient to use on daily basis when embedded into wearable devices in comparison to electrocardiography (ECG) [11]. The heart rate can be predicted at a lesser duration of time which is an advantage in the medical and health care sector to briskly determine any cardio-vascular problems. This paper exhibits the highest accuracy of 96.67% for Random Forest classifier when providing feature dataset as an input for classification. The future plan is to extend the proposed methodology in determining more number of vigorous high intensity activities by inclusion of other sensors like gyroscope to further increase the efficiency in prediction.

REFERENCES

- [1]. A.Reiss, I. Indlekofer, P. Schmidt and K. V. Laerhoven, "Deep PPG: Large-scale Heart Rate Estimation with Convolutional Neural Networks": MDPI Sensors, 19(14), 2019.
- [2]. G. Biagetti, P. Crippa, L. Falaschetti, L. Saraceni, A. Tiranti and C. Turchetti, "Dataset from PPG wireless sensor for activity monitoring", Elsevier, 2019, vol. 29.
- [3]. A.P. Psatha, A. Papaleonidas and L. Iliadis, "Machine Learning Modeling of Human Activity Using PPG Signals", ICCCI 2020: Lecture Notes in Computer Science, Springer, vol.12496, pp. 543-557, Nov. 2020.
- [4]. H. Gao, X.Wu, C. Shi, Q.Gao and J. Geng , "A LSTM-Based Realtime Signal Quality Assessment for Photoplethysmogram and Remote Photoplethysmogram", Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, pp. 3831-3840, 2021.
- [5]. M. Wójcikowski and B. Pankiewicz, "Photoplethysmographic Time-Domain Heart Rate Measurement Algorithm for Resource-Constrained Wearable Devices and its Implementation", Sensors (Basel), Mar.2020, 20(6): 1783.
- [6]. P. H. Charlton, P. Kyriacou, J. Mant, J. Alastruey, "Acquiring Wearable Photoplethysmography Data in Daily Life: The PPG Diary Pilot Study", 7th International Electronic Conference on Sensors and Applications session Wearable Sensors: MDPI, Nov. 2020.
- [7]. P. Casale, O. Pujol, and P. Radeva, "Human Activity Recognition from Accelerometer Data Using a Wearable Device", IbPRIA 2011: Pattern Recognition and Image Analysis, Lecture Notes in Computer Science, vol. 6669, pp. 289-296, Springer, 2011.
- [8]. M. Hossain, M.N. Sulaiman, "A Review on Evaluation Metrics for Data Classification Evaluations", International Journal of Data Mining & Knowledge Management Process (IJDKP), vol.5, no.2, pp.1-12, 2015.
- [9]. Z. Khan, A. Gul and A. Perperoglou , Ensemble of optimal trees, random forest and random projection ensemble classification", Advances in Data Analysis and Classification, Springer, vol. 14, pp. 97-116 , 2020.
- [10]. Jurek, Y. Bi, "A survey of commonly used ensemble-based classification techniques", The Knowledge Engineering Review, vol. 29:5, pp.551-581, 2013.
- [11]. R.G.R. Lekkala, S. Kuntamalla, "A novel approach for comparison of heart rate variability derived from synchronously measured electrocardiogram and photoplethysmogram", IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT), pp. 121-124, 2017.