Acceleration Plethysmogram Based Biometric Identification

Nur Azua Liyana Jaafar, Khairul Azami Sidek and Siti Nurfarah Ain Mohd Azam
Department of Electrical and Computer Engineering
Faculty of Engineering, International Islamic University Malaysia
53100 Jalan Gombak, Kuala Lumpur, Malaysia
azami@iium.edu.my

***Abstract*— This paper presents the feasibility study of Acceleration Plethysmogram (APG) based biometric identification system. APG signals are obtained from the second derivative of the Photoplethysmogram (PPG) signal. It has been reported from previous literature that APG signals contain more information as compared to the PPG signal. Thus, in this paper, the robustness and reliability of APG signal as a biometric recognition mechanism will be proven. APG signals of 10 subjects were acquired from the Multiparameter Intelligent Monitoring in Intensive Care II Waveform Database (MIMIC2WDB) which contains PPG signals with a sampling frequency of 125 Hz. The signals were later converted into an APG waveform. Then, discriminating features are extracted from the APG morphology. Finally, these APG samples were classified using commonly known classification techniques to identify individuals. Based on the experimentation results, APG signal when using Bayes Network gives an identification rate of 97.5 percentage as compared to PPG signal of 55 percentage for the same waveform. This outcome suggests the feasibility and robustness of APG signals as a biometric modality as compared to PPG signals.**

Keywords*— *APG signal; Biometric; Bayes Network; MIMIC2WDB; PPG signal.

I. INTRODUCTION

In today's world, people are becoming more insecure when it comes to giving their personal information to a third party. This is because they are afraid that other people might misuse their credentials for illegal purposes as today's crimes such as identity theft has increased rapidly in the past years. A survey conducted by UNISYS Security Index in 2014 showed that identity crime is an increasingly alarming issue across the globe [1]. The report reveals that 86 % of Mexicans are seriously concerned about identity theft whereas 78 % of Malaysians are very disturbed about unauthorized access to or misuse of their personal information. Additionally, in Australia and the United Kingdom, more than 51 % of people are worried about their personal credentials when involving any transactions which include personal credentials [1]. One of the key strategies to combat identity crime is to implement biometric system.

Recently, the usage of biological signals such as Photoplethysmogram (PPG) signals for the purpose of biometric recognition has been implemented [2, 3]. PPG signal has its own characteristics and features that portray a person's identity. However, the contour of the PPG waveform as shown in Fig 1 is very simple yet difficult to analyze and detect the phase changes of the contour.

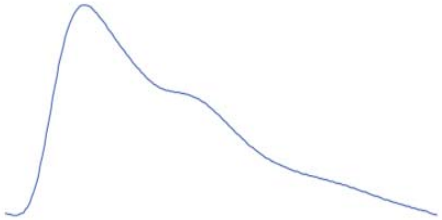

Fig 1: PPG waveform [4]

Hence, it requires other alternative methods that might assist the interpretation process of the PPG. Elgendi in [4] introduced the second derivative of PPG signal that may facilitate the analysis of the original PPG signal which is called acceleration plethysmogram (APG) signal as illustrated in Fig 2.

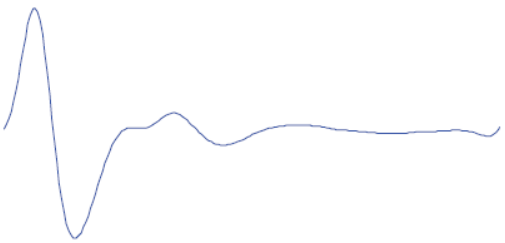

Fig 2: APG waveform [4]

APG is an indication of the acceleration of the blood in the finger. It has been reported from previous literature that APG signal contains more information such as visible phase change as compared to the PPG signal [5-9].

Singh and Nagpal explained the shape of APG signal consisting of seven different types where each type represents different condition of a person's heart [5]. Elgendi et. al. proposed the method of calculating the heart rate

variability using the detection of **a**-wave interval in APG waveform [6]. Nousou et. al. developed a diagnosis assistance system that uses APG signal acquired from the fingertip of a person to diagnose in a short period of time [7]. Rozi et. al. experimented the effects of exercise towards APG wave shapes [8]. Baek et. al. implemented second derivative of PPG to monitor the arterial condition [9].

However, previous researches on APG signal were focused on clinical purposes. Little has been said about the application of APG in identifying individuals. Thus, this study will be focusing on the implementation of APG signal for biometric identification purposes. The feasibility of the biometric system is investigated in this research.

The remaining of the paper is organized as follows; the next section describes the methodology of the study. Later, Section III elaborates about the classification performance of APG signal when applied to commonly used classification techniques. And finally in Section IV, the conclusion is layed out based on the experimentation and results in the previous section.

II. SYSTEM AND METHOD

The general architecture of our proposed system begins with signal acquisition of the PPG recordings, followed by the second differentiation of the PPG samples, then the pre-processing stage and later the extraction of APG signals as unique features by performing normalization and segmentation on the PPG data. Finally, these extracted features are applied to commonly used classifiers for identification purposes. The proposed identification system is summarized as in Fig 3 and these steps are briefly discussed in the next sub-sections.

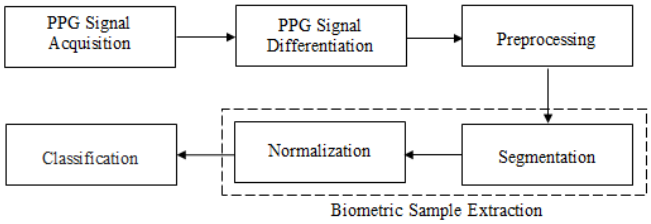

Fig. 3. The Proposed APG Based Biometric Identification System

A. Signal Acquisition

In this study, we acquired PPG signals taken from 10 different subjects. These recording were obtained from Physionet with duration of ten seconds each [10, 11]. These samples were fetched from the Multiparameter Intelligent Monitoring in Intensive Care II Waveform Database (MIMIC2WDB) which contains multiple physiologic signals collected from bedside patient monitors in adult and neonatal intensive care units (ICUs). These physiologic signals include PPG signals which were sampled at 125 Hz with 8-bit resolution.

B. Signal Differentiation

After collecting the PPG data, second order differentiation is performed to produce the APG signal. In the preliminary stage, the PPG data, which is denoted as x is differentiated to obtain the differentiated values denoted as y in Equation 1.

$$y(m) = x(n-1) - x(n) \tag{1}$$

where $m = 1, 2, 3, \ldots, (N-1)$ and $y(m)$ is the differentiated PPG signals.

$$\frac{\partial^2 y}{dx^2} = \frac{\partial}{\partial x}\left(\frac{\partial y}{dx}\right) \tag{2}$$

Next, the differentiated values are again differentiated to produce the second order derivative as in Equation 2 and construct the APG signal.

C. Pre-processing

The output of the differentiation process is converted to an APG signal. However, the signal is not clean and artifacts are prevalent in the waveform. Thus, a Butterworth low pass filter is used to mitigate the noise factor. Then, discriminant features are extracted to act as the input for the classifiers.

D. Biometric Sample Extraction

This stage is divided into two main steps which are segmentation and normalization.

i. Segmentation

Once APG signals have been acquired, it is segmented based on the amplitude characteristic of the wave. The procedure begins by identifying the **a** wave and making it the pivot since it represents to the highest and most obvious peak in an APG signal. From the **a** wave, equal numbers of data points are selected to the right and left of this reference. We repeat the previous steps in different time instances to collect more APG signals for every subject which would represent enrolment and recognition datasets.

ii. Normalization

After PPG samples have been segmented, normalization is applied to eradicate inherited artifacts due to baseline wanders from the PPG signal by leveling it to a common scale. By doing so, it would be much easier to analyze similarities of PPG signals. This technique is defined in Equation 3.

$$Normalization, N = \frac{x - \mu_x}{n_x} \tag{3}$$

where x is the PPG data, μ_x is the mean value of x and n_x is the data points in x. These normalized PPG samples then acts as the input for the classifiers.

E. Classification

Classification is a fundamental task in pattern recognition that involves the formation of a classifier. The induction of classification algorithms from data of pre-classified instances is a central issue in data mining. Various techniques to this problem are based on numerous functional representations such as decision trees, neural networks, decision graphs, and rules [12]. In this paper, we use two commonly applied classification method to identify the class labels for instances, each typically described by a set of features (attributes) for the APG signals. The classifiers are Bayes Network and k Nearest Neighbour. These induction algorithms are briefly discussed in the following sub-sections.

i. *Bayes Network*

In order to apply Bayesian Network as a classification algorithm, we compute $argmax_y$ P(y|**x**) using the probability distribution, P(U) denoted as in Equation 4

$$P(y \mid \mathbf{x}) = P(U) / P(\mathbf{x}) \; \alpha \; P(U) \tag{4}$$

where $y = x_0$, the class variable given a set of variables $\mathbf{x} = x_1, x_2, \ldots, x_k$ called APG attributes. A classification algorithm, h : $\mathbf{x} \rightarrow y$ is a function that maps instances of **x** to values of y.

Since all variables in x are identified, we calculate Equation 4 for all class values. The Bayesian Network first learns the network structure and the probability tables. Once a reliable network structure is known, the conditional probability tables for each of the variables can be approximated.

ii. *K Nearest Neighbour*

k Nearest Neighbour is an instance based learning algorithm. It defines the hypotheses from the training instances and is able to adhere its model to previously unrecorded data. It operates by obtaining the nearest neighbour or the majority class among *k* neighbours. In other words, it finds the most similar element to a given query element with similarity defined by the Euclidean distance. The algorithm is determined locally and all computation is deferred until the classification step. The aim of the algorithm is to determine the class of a new case based on the class of the *k* most similar database elements [13].

Assume vector $\langle x, f(x) \rangle$ where $f(x)$ is the function for each instance x which can be represented by the feature vector

$$\langle y_1(x), y_2(x), \ldots, y_n(x) \rangle$$

where $y_n(x)$ is the value of the nth attribute of instance x. The Euclidean distance is shown in Equation 5.

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n} (y_r(x_i) - y_r(x_j))^2} \tag{5}$$

where $d(x_i, x_j)$ is the distance between two instances, x_i and x_j. Then, find $\hat{f}(x_q)$ which is the value of the output of the classification algorithm as denoted in Equation 6.

$$\hat{f}(x_q) = \arg\max_{v \in V} \sum_{i=1}^{k} \delta(v, f(x_i)) \tag{6}$$

where $\delta(a,b) = 1$ if $a = b$ and where $\delta(a,b) = 0$. The value $\hat{f}(x_q)$ is also an approximate of $f(x_q)$ among k training dataset closest to x.

III. EXPERIMENTATION AND RESULTS

Figures 5 and 6 illustrates the identification process of subject 2 and 4 which are collected signals of two subject among approximately 13 500 ICU patients undergoing the first three stages of our proposed biometric recognition system as elaborated in Section II. As we can observe from the waveforms of the APG signals of these two subjects, it can be seen that the morphological wave shapes are distinguishable as compared to a PPG signal. It would be difficult to discriminate the peaks in a PPG signal specifically the diastolic peak as it would resemble the diastolic slope as shown in Fig 1. However, the systolic and diastolic peaks are prevalent and obvious in an APG signal as shown in Fig 4. The clear waveform obtained from the second derivation of PPG signal enable the visualization of individual features to be detected and analyzed easily.

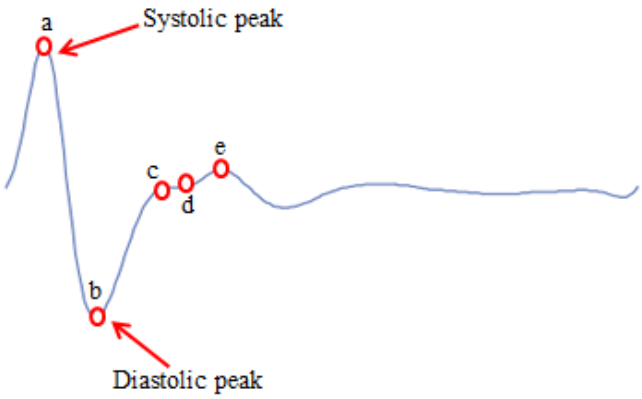

Fig. 4. APG waveform with distinguishable systolic and diastolic peaks consisting of **a**, **b**, **c**, **d** and **e** waves.

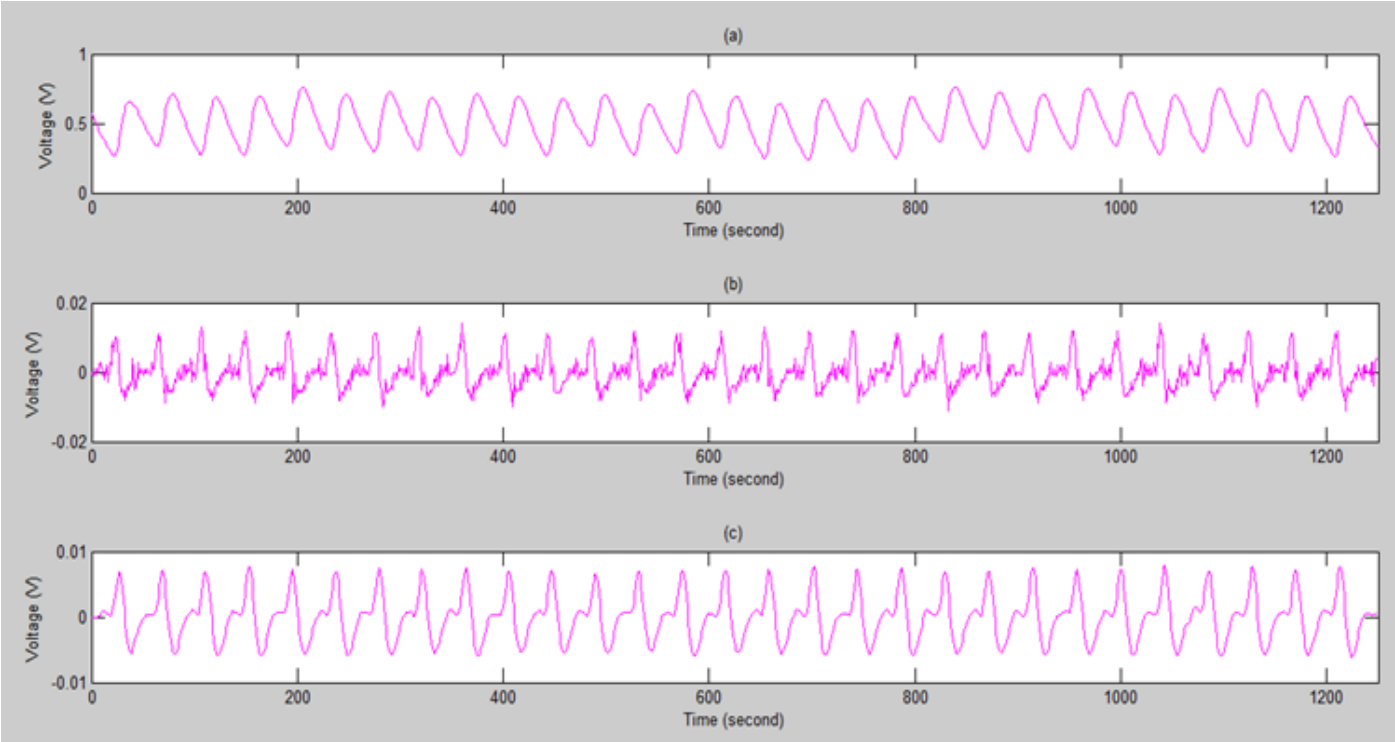

Fig. 5. (a) Raw PPG signal acquired from Subject 2 (b) Unfiltered APG signal and (c) Filtered APG signal of Subject 2.

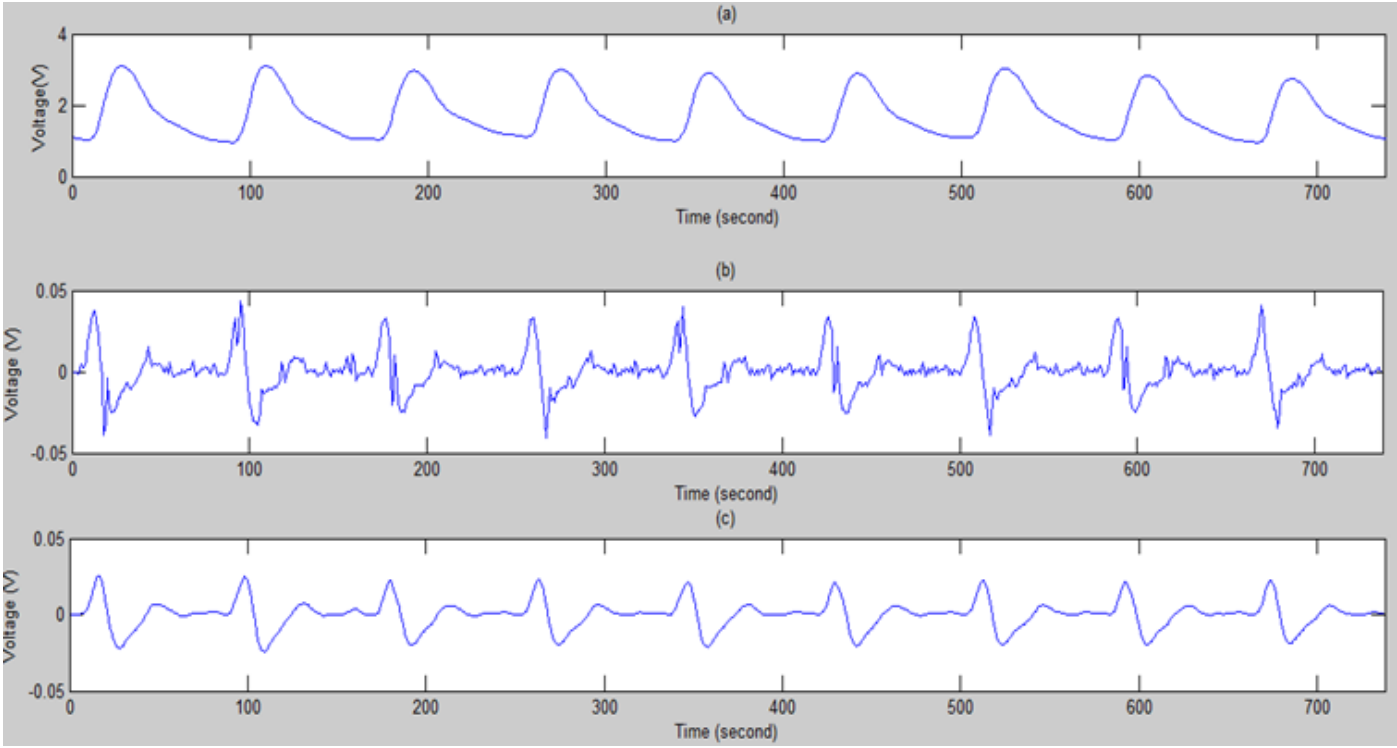

Fig. 6. (a) Raw PPG signal acquired from Subject 4 (b) Unfiltered APG signal and (c) Filtered APG signal of Subject 4.

978-1-4799-6879-4/15/$31.00 ©2015 IEEE

The waveform of APG signal consists of four systolic waves and one diastolic wave which are the **a**-wave (early systolic wave), **b**-wave (early systolic negative wave), **c**-wave (late systolic re-increasing wave), **d**-wave (late systolic re-decreasing wave) and **e**-wave (early diastolic positive wave). Each wave location reflects the closure of aortic valve and blood flow which can be used to monitor cardiac function.

The following stages of our proposed identification system are followed by the segmentation and normalization procedures. The implementation of these steps for subjects 2 and 4 are shown in figures 7 and 8. In these figures, different APG signals were acquired from varying time instances which were selected randomly.

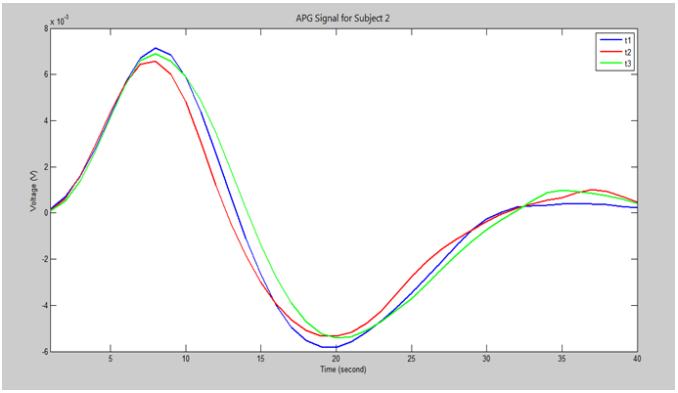

Fig 7: Subject 2's segmented and normalized APG signals

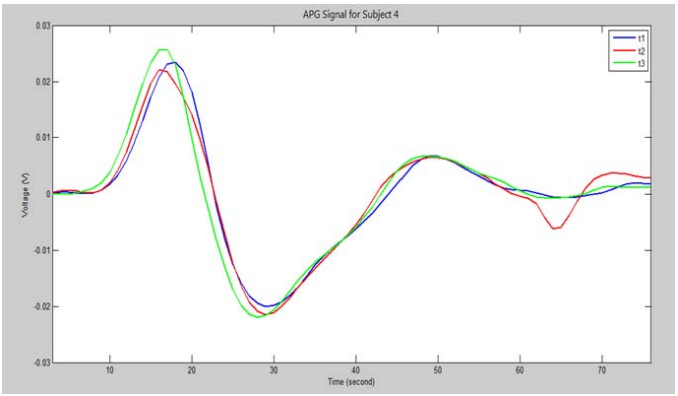

Fig 8: Subject 4's segmented and normalized APG signals

Recalling from Section II, we have tested our recognition approach by using 10 subjects from MIMIC2WDB. The APG for each subject comprises of 4 complete APG cycles, each representing 4 different instances. Therefore, the classification procedure was performed on 40 instances in total. Out of these 4 instances for a single subject, 2 instances are regarded as the enrolment samples whereas the remaining half is the recognition data. Later, a ten-fold cross validation method which assesses the generalization accuracy of the classification algorithms was implemented on the attributes to classify the subjects.

For our experimental analysis, we evaluated the classification accuracy as our performance metrics as it is a statistical measure on how efficient a classification algorithm correctly identifies individuals. In order to show the reliability of APG signals in comparison to PPG signals, we computed accuracy rates for both signals applying Bayes Network and k Nearest Neighbour. We used both signals (PPG and APG) from the same subject in the same time instance. For example, the PPG signals were obtained for a ten second duration. Thus, the APG signal was converted using the same ten second period from the same PPG signal as shown in figs 4 (a), 4(c), 5(a) and 5(c).

Based on the experimentation results, we found out that the classification accuracies of APG signal when applied to Bayes Network and k Nearest Neighbour outperformed the outcomes of PPG signal achieving identification rates of 97.5% and 90% as compared to 55% and 62.5% respectively. These classification accuracies are as summarized in Table 1. The output shows a dramatic increment of the classification accuracy when converted from PPG to APG waveform. These results support the capability of APG signal to display clear systolic and diastolic peaks in its morphological shape and as a results, produces better classification outcomes.

Furthermore, Bayes Network shows better classification accuracy as compared to k Nearest Neighbour due to three main reasons. One, Bayesian Network is capable of learning causal relationships, and thus can be implemented to understand about a problem domain and to predict the effect of its involvement. Two, Bayes Network has both a causal and probabilistic semantics where it is an ideal representation for the combination of prior knowledge and data. And three, Bayesian statistical methods in combination with Bayesian networks provides an effective and principled mechanism to avoid data over-fitting [27].

TABLE 1: CLASSIFICATION ACCURACIES WHEN COMPARING APG AND PPG SIGNALS APPLYING TWO CLASSIFIERS.

CLASSIFIER	CLASSIFICATION ACCURACY	
	APG SIGNAL	PPG SIGNAL
BAYES NETWORK	97.5 %	55 %
K NEAREST NEIGHBOUR	90 %	62.5 %

IV. CONCLUSION

As a conclusion, this study provides the proof-of-concept of applying APG signal for biometric identification purposes. Based on our knowledge, biometric recognition using APG signal has never been reported in previous literatures. The identification rate of 97.5 % as compared to 55 % proves the capability of APG signal to classify individuals better in

comparison to PPG signals. The outcome is significant as it suggests the feasibility and robustness of APG signals as a biometric modality. Additionally, APG based biometric identification is able to become a complementary mechanism for currently available biometric systems.

REFERENCES

[1] UNISYS Security Index, UNISYS, (Accessed on January 25th, 2015) http://www.unisyssecurityindex.com/usi/

[2] K. A. Sidek, I. Khalil, and H. F. Jelinek. "ECG Biometric with Abnormal Cardiac Conditions in Remote Monitoring System", *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1498-1509, November 2014.

[3] P. Spachos, G. Jiexin and D. Hatzinakos, "Feasibility study of photoplethysmographic signals for biometric identification", *2011 17th International Conference on Digital Signal Processing (DSP)*, pp.1-5, 6-8 July 2011.

[4] M. Elgendi. "On the Analysis of Fingertip Photoplethysmogram Signals", *Current Cardiology Reviews*, vol. 8, no. 1, pp. 14-25, 2012.

[5] M. Singh and S. Nagpal, "Features Extraction in Second Derivative of Finger PPG Signal: A Review." *IJCSC*, vol. 4, no. 2, pp. 1-5, September 2013.

[6] M. Elgendi, M. Jonkman and F. DeBoer, "Heart Rate Variability and the Acceleration Plethysmogram Signals Measured at Rest", *Biomedical Engineering Systems and Technologies: Communications in Computer and Information Science*, Springer Berlin Heidelberg, vol. 127, pp 266-277, 2011.

[7] N. Nousou, S. Urase, Y. Maniwa, K. Fujimura and Y. Fukui, "Classification of Acceleration Plethysmogram Using Self-Organizing Map," *International Symposium on Intelligent Signal Processing and Communications 2006 (ISPACS '06)*, pp. 681-684, 12-15 Dec. 2006.

[8] M. Rozi, M. A. M. Ali and M. B. I. Reaz, "Effects of exercise on the second derivative photoplethysmography (PPG) waveform," *2010 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, pp. 804-807, 6-9 Dec. 2010.

[9] H. J. Baek, J. S. Kim, Y. S. Kim, H. B. Lee, and K. S. Park, "Second Derivative of Photoplethysmography for Estimating Vascular Aging." *6th International Special Topic Conference on ITAB*, pp. 70-72, 2007.

[10] M. Saeed, M. Villarroel, A.T. Reisner, G. Clifford, L. Lehman, G.B. Moody, T. Heldt, T.H. Kyaw, B.E. Moody and R.G. Mark. Multiparameter intelligent monitoring in intensive care II (MIMIC-II): A public-access ICU database. *Critical Care Medicine*, vol. 39, no 5, pp. 952-960, May 2011.

[11] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff JM, P. Ch. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. K. Peng and H. E. Stanley. PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation*, vol.101,no.23,pp.215-220, 2000.

[12] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian Network Classifiers. *Machine Learning*, Springer, vol. 29, pp. 131-161, 1997.

[13] A G. Baptista and D. Silva, "How k-Nearest Neighbour Parameters Affects its Performance?", *38 JAIIO - Simposio Argentino de Inteligencia Artificial (ASAI 2009)*, pp. 95-106, 2009.