Machine learning application for classifying serum uric acid level with photoplethysmogram

Nazrul Anuar Nayan[1,2]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
[2]The Institute of Islam Hadhari
Universiti Kebangsaan Malaysia
Bangi, Malaysia
nazrul@ukm.edu.my

Hafifah Ab Hamid[1]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
Universiti Kebangsaan Malaysia
Bangi, Malaysia
p98362@siswa.ukm.edu.my

Mohd Zubir Suboh[1]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
Universiti Kebangsaan Malaysia
Bangi, Malaysia
p98999@siswa.ukm.edu.my

Nurliyana Muhammad Aminaldin[1]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
Universiti Kebangsaan Malaysia
Bangi, Malaysia
a160452@siswa.ukm.edu.my

Nurin Izzati Mohamad Azizul[1]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
Universiti Kebangsaan Malaysia
Bangi, Malaysia
a165324@siswa.ukm.edu.my

Mohamad Nazhan Mohd Nizar[1]
[1]Department of Electrical, Electronic & Systems Engineering, Faculty of Engineering & Built Environment
Universiti Kebangsaan Malaysia
Bangi, Malaysia
a165519@siswa.ukm.edu.my

2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES) | 978-1-7281-4245-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/IECBES48179.2021.9398763

***Abstract*— Cardiovascular disease (CVD) is a cause of 17.9 million deaths each year globally. Among the risk factors of CVD are hypertension, hyperlipidemia, hyperglycemia, and hyperuricemia. In hyperuricemia, serum uric acid (SUA) level can be detected only with a blood test, a microinvasive procedure that is costly and incur minimal pain and requires patients to travel to a hospital for regular monitoring. In this paper, a machine learning technique is proposed for SUA level classification. A total of 74 subjects participated in this study, and 20 features were extracted from the pre-processed photoplethysmogram (PPG) morphology. Blood tests were performed, and the results were used as inputs and labeled accordingly. Results show that based on p-value, the onset-diastolic width of the right and onset-systolic width of the left hand are the most significant features for SUA classification. Artificial neural network has an accuracy of 86.36%, which is higher than that of three machine learning methods: k-nearest neighbor, support vector machine, and decision tree.**

Keywords— CVD, PPG, hyperuricemia, machine learning, artificial neural network

I. INTRODUCTION

Cardiovascular diseases (CVDs) caused approximately 17.9 million deaths in 2016, representing 31% of all global deaths, according to the WHO's official website in 2017. In Malaysia, 15% of the 109,164 medically certified death in 2019 is caused by ischaemic heart diseases, ranking it the principal cause of death [1]. The risk factors of CVDs are hyperglycemia, hyperlipidemia [2], hypertension, and hyperuricemia [3]. Hyperuricemia is a condition wherein serum uric acid (SUA) level is more than 416 µmol/L [4]. It occurs when the body produces excessive uric acid (UA) or when the body cannot excrete sufficient UA and maintain the required level. UA level is highly affected by the intake of purine-rich foods and beverages, including alcoholic drinks, soft drinks, red meat, internal organs of edible animals, some types of seafood, and beans. SUA level can be detected only through a blood test, which is a microinvasive procedure. Excess SUA can form crystals around the joints and kidneys. When white blood cells attack the crystals, inflammation and pain occurs [5]. Inflammation can lead to blood clots, which in turn contribute to the onset of heart diseases. Statistics show that hyperuricemia can cause elevated high blood pressure, which is one of the major factors of CVD. Therefore, by continuously monitoring SUA level, CVD can be predicted, and treatment can be performed early. Thus, probability of recovery increases, and the risk of death from CVD is reduced.

Photoplethysmogram (PPG) signals are easy to obtain. They contain the optically detected information of blood volume changes in blood vessel raw PPG signals. However, they usually contain noises, such as baseline wandering and power line interference. It is essential to remove noises from PPG signal before proceeding for further works [6].

Machine learning (ML) approaches have been used in SUA prediction on the basis of basic health checkup results, especially blood test, which is an invasive method. Ichikawa [7] and Lee [8] used ML in predicting hyperuricemia in patients. Gradient-boosting decision tree, random forest, and logistic regression show similar predictive capability in identifying patient at a high risk of hyperuricemia [7]. According to a previous study [8], naïve Bayes ML shows the highest balanced classification rate for predicting hyperuricemia status. Recent publication on predicting UA used boosted decision tree regression models and showed promising results in Bangladesh populations [9]. However, SUA level prediction using ML from PPG data, which is one of the promising method in assessing health condition now, is rarely conducted.

In this paper, SUA level is noninvasively classified with PPG data. This research has been approved by the Research and Ethics Committee of the Hospital Universiti Kebangsaan Malaysia, with registration number UKM.PPI.800-1/3/21.

II. METHODOLOGY

The methodology is categorized into five main parts: data acquisition, signal preprocessing, feature extraction, statistical analysis, and machine learning algorithm development for the classification, as shown in Fig. 1.

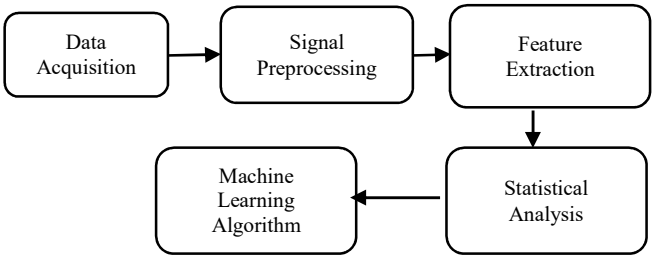

Fig. 1. Methodology.

A. Data Acquisition

A single data sample consisted of two PPG signals, blood test result, medical history, and personal information. A total of 74 (42 males, 32 females) nonsmoking subjects participated in this study. The subjects were 19–80 years old (mean, 39; median, 37).

PPG signals were obtained from both hands of each subject with pulse oximeter, which was connected to a signal recording software installed in a personal computer and attached to the tip of the index finger of each subject. The duration of PPG signal recording was 10 mins for each hand. This procedure allowed the acquisition of sufficient amount of useful information from PPG. Apart from PPG signal detection, blood tests were used in obtaining SUA data. The subjects were divided into "low" and "high" according to the mean SUA level.

B. Signal Preprocessing

Signal preprocessing includes algorithms development for signal quality indices (SQI), baseline and high frequency removal, and four fiducial peaks. MATLAB was used for the entire development process.

i. Signal quality indices

Reliable raw PPG signals were obtained using SQI. The PPG signal extrapolated from the 10 s sample was between 40 and 180 beats per minute. PPG pulse peak gap must not exceed 3 s in order that not more than one beat is missed. For the ratio of the maximum beat-to beat interval to the minimum beat-to-beat interval within sample should be less than 2.2 [10].

ii. Baseline and High Frequency Removal

Raw PPG signal was filtered with a 4th order Bessel bandpass filter at cutoff frequencies at 0.8 and 8 Hz. Signals below 0.8 Hz would be motion artefacts shown by baseline wandering [11]. After the low frequencies were filtered, the PPG signal contained decreased amounts of DC noises causing baseline wandering. Signals above 8 Hz might be powerline interferences. We observed that the filtered signal had an amplitude offset. An auto-offsetting algorithm was used in overcoming the problem. The algorithm simply determined whether or not a y-axis value was below zero and brought the negative value back to a positive value by offsetting the signal by the difference between zero-amplitude and the largest negative value detected.

iii. Peak Detection

The four fiducial points from PPG signals are systolic peak, dicrotic notch, diastolic peak, and pulse onset, as shown in Fig. 2. The Pulse Wave Delineator, a function that has been developed by ABing, was used in detecting the points [12].

C. Feature Extraction

A total of 16 features were extracted from two PPG pulses of each hand, that is. the time interval or the width between the points (unit, s), as shown in Fig. 2. Another four features were in the amplitude domain (unit: mm), as demonstrated in Fig. 3. The prefixes O, S, N, and D were used in describing onset, systolic peak, dicrotic notch, and diastolic peak, respectively, whereas the number 1 or 2 (e.g., O1, S1, N2, and D2) indicated features for the next pulse. Thus, the total of extracted features in this study was 20 for each hand.

A feature was not represented by a single value. It contained a set of data. For example, time differences between onsets formed a feature of "onset-to-onset." Looping is the technique usually used in obtaining these data. If there are n number of onsets, the n-1 number of onset-to-onset time difference can be extracted. Thus, when n pairs of two points of interest are present, the n-1 number of loops is needed. For each PPG signal, an array of values is present in a feature. Then, the median of the values in each feature is obtained. Median is preferred over mode or mean because it is more resilient to outliers.

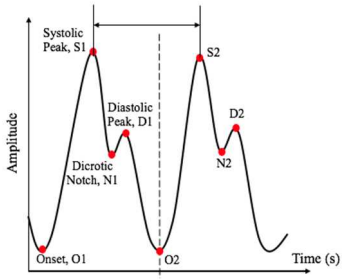

Fig. 2. Fiducial points and time domain features of two PPG pulses.

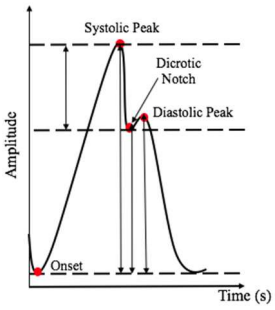

Fig. 3. Fiducial points and features of amplitude domain.

D. Statistical Analysis

The main purpose of this research is to predict SUA level classes (low or high level) with 40 extracted features from PPG signals taken from the right and left hands. The analysis of mean for each feature were performed using independent sample t-test with confidence interval of 95%. The mean, standard deviation, and p-values were observed for the identification of the most significant features that can distinguish both classes. Features were ranked according to their p-values, and highly ranked significant features were used as inputs for the classifiers.

E. Machine Learning

Four machine learning techniques were used in classifying the selected features into two groups: Control and Case for low and high SUA levels, respectively. The values used in dividing the groups were the "mean" SUA values for each male and female taken with the gold standard SUA level test.

The data were then randomly divided into train and test data in a ratio of 70:30. An artificial neural network (ANN),

k-nearest neighborhood (kNN), support vector machine (SVM), and decision tree (DT) were selected in this study. Each classifier was trained with different number of selected features by using a fivefold cross validation method. The specificity (Sp), sensitivity (Se), and accuracy (Acc) of each model were determined and used in assessing the performance of each classifier. The model with the highest Acc and minimal number of features was regarded as the best classification model.

For ANN, three layers of feedforward network was set with varying number of input neurons (based on the number of features ranging from 1 to 17), different number of hidden neurons at each layer (from 1 neuron to 30 neurons), and one output neuron for low and high SUA levels. The network was trained using the Levernberg-Marquadt training algorithm and "log-sigmoid" transfer functions at the first and second hidden layers. The "purelin" transfer function at the output layer was also used. The training goal was set at 1×10^{-7} with maximum epochs of 1000. Other parameters were set by default with MATLAB Classification Leaner App. Multiple kernels were used in the SVM for the creation of a hyperplane that can separate the groups. kNN was set with different number of neighbors by using Euclidean distance. Gini's diversity index (gdi) was used with different number of splits in the DT classifier.

III. RESULT AND DISCUSSION

A. Data Acquisition

Fig. 4 shows the image taken during the acquisition of PPG from a subject. A pulse oximeter was attached to a subject's fingertip to capture PPG signal. An arteriograph was used to record arterial blood pressure noninvasively. An example of a 6 min raw PPG signal acquired with 100 Hz sampling rate is shown in Fig. 5.

Fig. 4. Data acquisition using a pulse oximeter and arteriograph.

The UA level obtained from blood test results of the male subjects ranged from 230.0 to 651.2 µmol/L and those obtained from the female subjects ranged from 170.0 to 452.0 µmol/L. Thus, the mean or classification values for this study were 460 and 290 µmol/L for men and women, respectively. Subjects with SUA levels higher than the classification values were included in the "high" class, and those with levels lower than the classification values were included in the "low" class. Table I shows all 74 subjects with their respective classes.

TABLE I. SUBJECT AND CLASS DISTRIBUTION

	LOW SUA<460 (Male) SUA<290 (Female)	HIGH SUA≥460 (Male) SUA≥290 (Female)	Total
Male	20	22	**42**
Female	17	15	**32**
Total	**37**	**37**	**74**

B. Signal Preprocessing

i. SQI

Reliable PPG signals were obtained by using SQI. This process selects good signals for further analysis. SQI was used to reduce errors in the extraction of features from the PPG signals. Fig. 5 shows how SQI algorithm processes PPG signal data of a subject. The signals highlighted in blue had SQI of 0.95 and above. They consisted of 81% of the total recorded signals. The feature extraction algorithm selected high SQI signals and removed low ones.

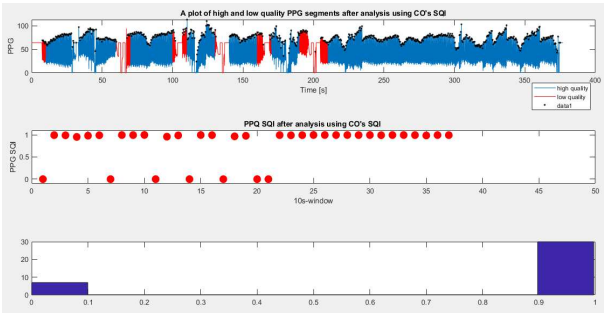

Fig. 5. Raw PPG signal with SQI output result.

ii. Baseline wandering and high frequency removal

Raw PPG signals were filtered with the 4th order Bessel bandpass filter with pass band at 0.8 and 8 Hz and auto-offset. The process used in comparing unprocessed and processed PPG signals is illustrated in Fig. 6.

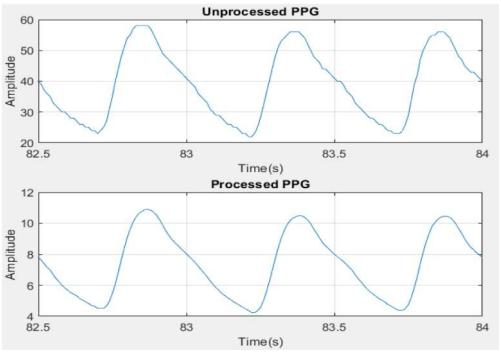

Fig. 6. Raw and preprocessed signals.

iii. Peak Detection

Pulse onsets, systolic peaks, dicrotic notches, and diastolic peaks present in PPG signals were detected using a delineator function. PPG signal with fiducial points detected is shown in Fig. 7. After all the fiducial points were detected, all the values of 20 features were extracted using windowing and looping methods for every PPG signal. For each PPG signal, the

median of the values in the feature was obtained and used as the representation value of each feature.

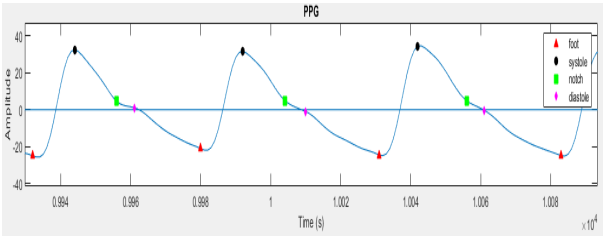

Fig. 7. The extracted four fiducial points.

C. Feature Extraction

As mentioned in the previous section, the features from left- and right-hand PPG were combined. Hence, a total of 40 features were obtained from each subject.

D. Statistical Study

The statistical values of all 40 features are listed in Table II. The underscore sign as in RO1_D1 depicted a time domain feature. The features were sorted in a way that the smallest p-value was at the top. Only two features were significant: RO1_D1 (onset-diastolic peak of the right hand) and LO1_S1 (onset-systolic peak width of the left hand) because their p-values were equal to or lower than 0.05. The prefix "R" represents the right hand, and "L" represents the left hand. However, more features are required to obtain a more accurate classification model. In this study, the top 17 features with p-values lower than 0.5 were selected for evaluation by ML techniques. Thirteen of these features were obtained from the left arm signals.

E. Machine Learning

Each classifier was trained with different number of selected features (1–17 features) as the inputs. A series of training was conducted for each input. The setting used was described in the ML subsection in Methodology. The best accuracy obtained for every input in each classification model are recorded in Table III. The Acc trend is illustrated in Fig. 8.

The ANN classifier showed slight increment in its Acc, showing an increase when at least 13 features were used. The ANN had the highest test Acc of 86.36% when eight features with nine hidden neurons at the first hidden layer and 14 neurons at the second hidden layer were used. The network provided acceptable percentage of Sp of 81.82% and good Sn of 90.91%. The Acc of the SVM classifier was slightly lower than that of the ANN (81.82%) when 12 features were used. The kNN classifier was able to differentiate the low and high classes at only 63.64% Acc, and the DT classifier showed inconsistently low Acc with different number of features. However, both the kNN and DT classifiers suggested that the last five features were reducing the Acc.

TABLE II. STATISTICAL VALUE FOR ALL 40 FEATURES

No	Feature	Control Group (Low SUA level)		Case Group (High SUA level)		p-value
		Mean	Std. Deviation	Mean	Std. Deviation	
1	RO1_D1	23.811	4.453	21.892	3.357	0.040
2	LO1_S1	34.500	4.640	32.108	5.640	0.050
3	RO1_N1	24.865	5.421	22.741	5.498	0.099
4	LO1_N1	24.189	4.360	22.432	4.896	0.107
5	RO1S1	0.1427	0.020	1.823	7.151	0.157
6	LD1O2	0.212	0.055	0.228	0.057	0.219
7	RS1_N1	14.770	3.435	13.865	3.483	0.264
8	LS1O2	0.371	0.059	0.388	0.067	0.270
9	LN1S2	0.376	0.049	0.391	0.072	0.277
10	LO1_D1	22.581	4.789	21.486	3.922	0.286
11	LS1D1	0.171	0.025	0.179	0.039	0.307
12	LS1_N1	14.811	3.560	13.851	4.920	0.340
13	LO1O2	0.517	0.067	0.532	0.082	0.393
14	LS1S2	0.517	0.067	0.532	0.083	0.396
15	LD1S2	0.358	0.065	0.372	0.075	0.418
16	LN1O2	0.239	0.053	0.249	0.057	0.440
17	LD1D2	0.517	0.066	0.531	0.081	0.493
18	LD1N2	0.370	0.167	0.394	0.172	0.541
19	RS1D1	0.180	0.033	0.1745	0.037	0.542
20	RO1_S1	32.716	5.797	32.027	3.734	0.545
21	RD1O2	0.222	0.058	0.229	0.057	0.566
22	LN1N2	0.522	0.068	0.531	0.079	0.600
23	RO1D1	0.302	0.020	0.306	0.044	0.632
24	RS1O2	0.380	0.061	0.387	0.068	0.647
25	RD1S2	0.365	0.067	0.372	0.076	0.667
26	RS1S2	0.523	0.067	0.530	0.085	0.682
27	LO1S1	0.145	0.024	0.143	0.023	0.710
28	RD1D2	0.5235	0.065	0.530	0.085	0.725

29	RN1O2	0.247	0.062	0.251	0.060	0.747
30	LO1D1	0.301	0.036	0.303	0.032	0.757
31	RN1S2	0.389	0.060	0.394	0.069	0.762
32	RN1N2	0.530	0.062	0.534	0.084	0.795
33	LN1D1	0.111	0.149	0.115	0.156	0.891
34	RO1N1	0.279	0.056	0.278	0.056	0.901
35	RN1D1	0.116	0.156	0.113	0.142	0.917
36	LS1N1	0.132	0.032	0.131	0.038	0.922
37	LO1N1	0.276	0.052	0.275	0.055	0.922
38	RD1N2	0.387	0.166	0.340	0.169	0.942
39	RS1N1	0.136	0.039	0.136	0.039	0.976
40	RO1O2	0.524	0.066	0.523	0.095	0.977

TABLE III. TEST PERFORMANCE FOR FOUR DIFFERENT CLASSIFIERS WITH DIFFERENT NUMBER OF FEATURES

No. of features	ANN			DT			SVM			kNN		
	Sp, (%)	Se, (%)	Acc, (%)	Sp, (%)	Se, (%)	Acc, (%)	Sp, (%)	Se, (%)	Acc, (%)	Sp, (%)	Se, (%)	Acc, (%)
1	54.55	90.91	72.73	54.55	90.91	72.73	45.45	63.64	54.55	81.82	18.18	50.00
2	81.82	81.82	81.82	45.45	81.82	63.64	63.64	63.64	63.64	54.55	72.73	63.64
3	54.55	100	77.27	81.82	45.45	63.64	36.36	54.55	45.45	81.82	45.45	63.64
4	90.91	63.64	77.27	81.82	45.45	63.64	45.45	63.64	54.55	63.64	36.36	50.00
5	63.64	90.91	77.27	72.73	54.55	63.64	36.36	45.45	40.91	100	0.00	50.00
6	63.64	100	81.82	36.36	90.91	63.64	45.45	36.36	40.91	63.64	63.64	63.64
7	90.91	72.73	81.82	54.55	81.82	68.18	72.73	72.73	72.73	63.64	45.45	54.55
8	81.82	90.91	86.36	72.73	36.36	54.55	45.45	72.73	59.09	36.36	54.55	45.45
9	72.73	90.91	81.82	63.64	63.64	63.64	63.64	63.64	63.64	54.55	45.45	50.00
10	81.82	90.91	86.36	63.64	63.64	63.64	72.73	81.82	77.27	45.45	54.55	50.00
11	81.82	90.91	86.36	63.64	63.64	63.64	54.55	54.55	54.55	54.55	63.64	59.09
12	72.73	100	86.36	72.73	72.73	72.73	72.73	90.91	81.82	54.55	63.64	59.09
13	**90.91**	**81.82**	**86.36**	54.55	18.18	36.36	54.55	72.73	63.64	54.55	63.64	59.09
14	81.82	81.82	81.82	54.55	18.18	36.36	72.73	90.91	81.82	45.45	63.64	54.55
15	81.82	81.82	81.82	54.55	18.18	36.36	72.73	90.91	81.82	27.27	63.64	45.45
16	72.73	90.91	81.82	63.64	27.27	45.45	63.64	72.73	68.18	36.36	72.73	54.55
17	81.82	81.82	81.82	36.36	72.73	54.55	63.64	90.91	77.27	63.64	27.27	45.45

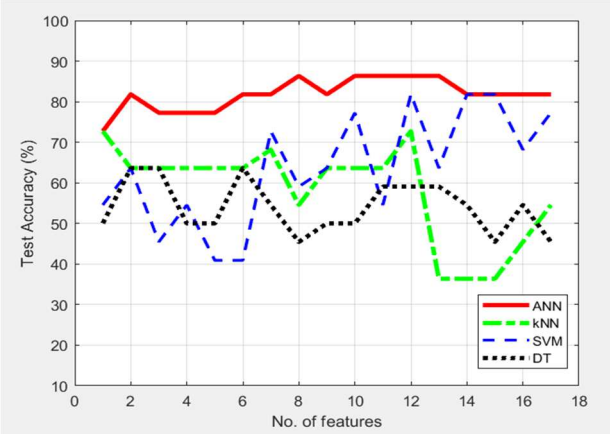

Fig. 8. Test Acc trend with different number of features.

IV. CONCLUSION

PPG data were collected from 74 participants in this study. Then, 40 features were extracted from each subject. Two features were significant, namely, the "onset-diastolic-peak" of the right hand and "onset-systolic-peak" of the left hand, as their p-values were equal to or lower than 0.05. For the ANN, seven models with different hidden layer configurations achieved training and testing Acc of over 80%, which was the highest Acc achieved with the traditional ML algorithm. In short, the ANN was able to increase the prediction Acc for the UA level groups. In building an ML model that is more accurate and consistent in performing classification, more data samples are needed. In a nutshell, the classification of SUA level was achieved with ML techniques. In the future, the ML model trained in this study can be converted into smart wristband applications, and everyone would be able to monitor their SUA levels anytime and anywhere.

ACKNOWLEDGMENT

The authors would like to thank the Ministry of Education, Malaysia and Universiti Kebangsaan Malaysia for supporting this research through research grant FRGS/1/2019/TK04/UKM/02/4.

REFERENCES

[1] Department of Statistics of Malaysia, "Statistics on Causes of Death, Malaysia," Nov 2020.

[2] V. Ormazabal, S. Nair, O. Elfeky, C. Aguayo, C. Salomon, and F. A. Zuñiga, "Association between insulin resistance and the development of cardiovascular disease," Cardiovascular diabetology, vol. 17, no. 1, p.122, 2018.

[3] M. Kuwabara, "Hyperuricemia, cardiovascular disease, and hypertension," Pulse, vol. 3, no. 3-4, pp.242-252, 2015.

[4] R. Terkeltaub, and N. L. Edwards, "Gout : Diagnosis and Management of Gouty Arthritis and Hyperuricemia," Professional Communication, Inc., vol. 4, 2016.

[5] S. Y. Morris, "Hyperuricemia: Symptoms, Treatment, and More – Healthline" https://www.healthline.com/health/hyperuricemia, 2018.

[6] T. D. Daniel, M. N. Ungureamu, and R. Strungaru, "Cancelling Harmonic Power Line Interference in Biopotentials," Compendium of New Technique in Harmonic Analysis, Intech, 2018.

[7] D. Ichikawa, T. Saito, W. Ujita, and H. Oyama, "How can machine-learning methods assist in virtual screening for hyperuricemia? A healthcare machine-learning approach," Journal of biomedical informatics, vol. 64, pp.20-24, 2016.

[8] S. Lee, E. K. Choe, and B. Park, "Exploration of machine learning for hyperuricemia prediction models based on basic health checkup tests," Journal of clinical medicine, vol. 8, no. 2, p.172, 2019.

[9] M. B. Sampa, M. N. Hossain, M. R. Hoque, R. Islam, F. Yokota, M. Nishikitani, and A. Ahmed, "Blood Uric Acid Prediction With Machine Learning: Model Development and Performance Comparison," JMIR medical informatics, vol. 8, no. 10, p.e18331, 2020.

[10] C. Orphanidou, T. Bonnici, and P. Charlton, "Signal -Quality Indices for the Electrocardiogram and Photoplethysmogram: Derivation and Applications to Wireless Monitoring," vol. 19, no 3, 2015.

[11] G. Slapnicar, N. Mlakar, and M. Lustrek, "Blood Pressure Estimation from Photoplethysmogram using a spectrotemporal deep neural network," Sensors, vol. 19, no. 15, 2019.

[12] ABing, "Pulse Waveform Delineator," https://www.mathworks.com/matlabcentral/fileexchange/29484-pulse-waveform-delineator, MATLAB Central File Exchange, Retrieved December 28, 2020.

[13] A. S. Meidert, and B. Saugel, "Technique for non-invansive monitoring of arterial blood pressure," Frontiers in Medicine, vol. 4, pp.1-6, 2017.

[14] Z. Taha, L. Shirley, M. M. A. Razman, "A review on non-invasive on non-invasive hypertension monitoring system by using photoplethysmography method," Movement, Health & Exercise, vol. 6, no. 1, pp.47-57, 2017.