

Article

Identification of Coronary Artery Diseases Using Photoplethysmography Signals and Practical Feature Selection Process

Amjed S. Al Fahoum ^{1,*} , Ansam Omar Abu Al-Haija ^{1,2} and Hussam A. Alshraideh ²¹ Biomedical Systems and Informatics Engineering Department, Yarmouk University, Irbid 21163, Jordan² Industrial Engineering Department, JUST, Irbid 22110, Jordan

* Correspondence: afahoum@yu.edu.jo

Abstract: A low-cost, fast, dependable, repeatable, non-invasive, portable, and simple-to-use vascular screening tool for coronary artery diseases (CADs) is preferred. Photoplethysmography (PPG), a low-cost optical pulse wave technology, is one method with this potential. PPG signals come from changes in the amount of blood in the microvascular bed of tissue. Therefore, these signals can be used to figure out anomalies within the cardiovascular system. This work shows how to use PPG signals and feature selection-based classifiers to identify cardiorespiratory disorders based on the extraction of time-domain features. Data were collected from 360 healthy and cardiovascular disease patients. For analysis and identification, five types of cardiovascular disorders were considered. The categories of cardiovascular diseases were identified using a two-stage classification process. The first stage was utilized to differentiate between healthy and unhealthy subjects. Subjects who were found to be abnormal were then entered into the second stage classifier, which was used to determine the type of the disease. Seven different classifiers were employed to classify the dataset. Based on the subset of features found by the classifier, the Naïve Bayes classifier obtained the best test accuracy, with 94.44% for the first stage and 89.37% for the second stage. The results of this study show how vital the PPG signal is. Many time-domain parts of the PPG signal can be easily extracted and analyzed to find out if there are problems with the heart. The results were accurate and precise enough that they did not need to be looked at or analyzed further. The PPG classifier built on a simple microcontroller will work better than more expensive ones and will not make the patient nervous.

Keywords: PPG signals; detection applications; cardiorespiratory function; feature combination; biomedical analysis; biomedical signal processing



Citation: Al Fahoum, A.S.; Abu Al-Haija, A.O.; Alshraideh, H.A. Identification of Coronary Artery Diseases Using

Photoplethysmography Signals and Practical Feature Selection Process.

Bioengineering **2023**, *10*, 249.

<https://doi.org/10.3390/bioengineering10020249>

<https://doi.org/10.3390/bioengineering10020249>

Academic Editors: Antonio Brunetti and Domenico Buongiorno

Received: 9 January 2023

Revised: 19 January 2023

Accepted: 7 February 2023

Published: 13 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Cardiovascular diseases (CVDs) are illnesses of the heart and blood vessels that result in inadequate arterial blood flow to many essential organs [1]. Recent data worldwide shows that the number of people with heart disease is increasing. CVDs were responsible for about 30% of deaths worldwide in 2018 and about 36% of deaths in Jordan in 2020, according to the Jordanian Ministry of Health (MOH) [2]. So, sensitive, practical, and cost-effective ways to monitor CVDs early on are needed, in order to speed up efforts to stop the rise of CVD-related deaths [3]. Classification algorithms based on machine learning are commonly employed to anticipate and categorize the many features of biomedical signals and molecules, such as cardiac anomalies, toxicity, or biological activity [4–8]. Three datasets were generated to test different classification models based on machine learning [4]. Further, they applied many classification performance indicators from many fields. To compare the two groups, they used a novel method that was based on traditional chemometric techniques, such as the sum of ranking differences (SRD) and analysis of variance (ANOVA) [8].

Risk factors, such as smoking, obesity, high blood pressure, family history, and others, can be used to predict all cardiovascular diseases [9]. On the other hand, the goal of the healthcare revolution was to make monitoring technologies, such as PPG, that were small, cheap, and easy to use and could predict the start of CVDs. PPG is a non-invasive method of examining a specific body signal that provides detailed information about the heart and lungs. Several researchers focused on finding links between the PPG signal and events in the heart and blood vessels, so they could predict CVDs [9–12]. They did this by relying on extracting the PPG signal's temporal domain properties [9–12]. Aboy et al. presented a pressure-based technique for automatic beat detection [13]. Solosenko et al. automatically exploited PPG signals to detect pre-ventricular contraction (PVC) [14]. Yousefi et al. [15] found a way to find PVC automatically by combining the chaotic nature of the PPG signal with higher-order statistics (HOS). They used several chaotic and statistical features, such as the Lyapunov exponent, skewness, kurtosis, fuzzy entropy, and spectral entropy, which were taken from the signals. Principal component analysis (PCA) was used to figure out which data points should be used to group the data. Poloania et al. and Abu Elhijja et al. developed algorithms for classifying cardiac arrhythmias based on PPG [16,17]. By taking PPG and ECG measurements, Nano et al. figured out the time between heartbeats and pulses [18]. Quality assessment for signal reliability estimation was proposed in [19] for PPG beat recognition and morphology estimation. Recently, Prabhakar et al. [20] developed a few metaheuristic techniques for reducing dimensionality. Post-classifiers then used the reduced values to separate PPG normal and CVD signals.

Recent strategies have focused on figuring out the best set of PPG characteristics that can be used to classify things helpfully. Data mining techniques, which are used to obtain useful information from data, can be used to find and emphasize differences in PPG time-domain characteristics between healthy human subjects and CVD patients [21]. Al-Fahoum et al. used data mining and a signal processing strategy called “multiple signal classification” to classify the difference between CVD patients and healthy normal subjects [12]. The primary applications of data mining techniques are the classification, prediction, and grouping of observations [22,23]. Improving the performance and usefulness of data mining algorithms means finding the best way to pick the most important attributes [15,24–28]. Recent research publications [28–33] demonstrated the need for automatic detection strategies for identifying arrhythmias using cutting-edge PPG-based techniques. The results from [28] to [33] supported the study's goal. Based on an analysis of the time-domain features of the PPG signal, this study aims to find the best set of studied features that use data mining techniques to combine the most critical features of the PPG signal with other demographic information to tell the difference between healthy human subjects and CVD patients.

2. Materials and Methods

2.1. Data Collection

The primary goal of this study is to distinguish between normal and abnormal cardiovascular cases by combining time-domain features of the PPG signal with a set of subject demographic characteristics. Data from 360 people (200 healthy and 160 with CVD) were collected for solid results. As for the PPG, data collection from CVD patients was carried out at the internal medicine clinic of Princess Basma Hospital. The study focused on five common types of CVDs: acute coronary syndrome (ACS) (63 patients), cerebrovascular accident (CVA) (50 patients), deep vein thrombosis (DVT) (23 patients), heart failure (HF), and atrial fibrillation (AF) (13 patients each). A chart of the five CVD cases that were studied, each with its percentage from the CVD sample, is displayed in Figure 1.

This study employed the PO-80 pulse oximeter from Beurer Healthcare in Germany to take non-invasive PPGs (pulse rate), SpO₂, and heart rate measurements. PO-80 is a portable, rechargeable, small (L 57 mm × W 32 mm × H 30 mm), lightweight (about 42 g), and compact device that comes with “SpO₂-Viewer/Manager” software that works with Windows. Using a USB connection, the software can send measured data from the device to

the PC, which can be viewed in real-time. The sensor has an accuracy of $\pm 2\%$ for SpO_2 in the range of 30–100% measurement and ± 2 beats/minute for pulse rate monitoring in the range of 30–250 beats/minute, according to its specifications.

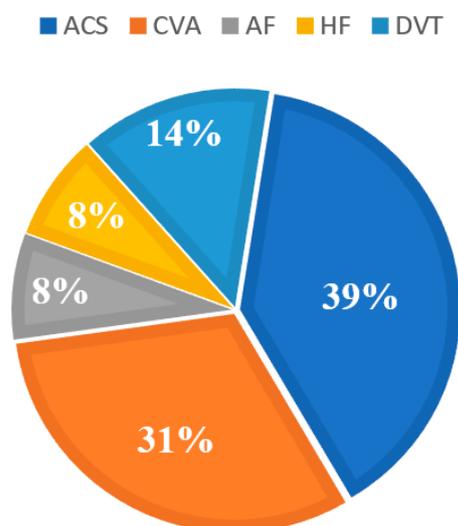


Figure 1. A chart of the five CVDs cases.

The SpO_2 sensor comprises an emitter and a detector housed in a finger clip probe. The emitter comprises a group of light-emitting diodes (LEDs) that give off red light at 660 nm and infrared light at 905 nm. The detector, in contrast, is a silicon-based photodiode. The method used to measure PPG is conducted by inserting one finger into the finger opening of the pulse oximeter. By pressing the function button on the PO-80 sensor, the oxygen saturation of the blood's hemoglobin and the heart rate can be measured non-invasively [34].

2.2. Data Description

In addition to the PPG signal recorded for at least 2 min while the volunteer was at rest, the following ten variables about the volunteer's background and health were also collected.

2.2.1. Demographic and Health Status Variables

Age

Age is considered a primary factor highly correlated with arterial stiffness, and its effect is reflected in the shape of the PPG signal [23,24].

Height, Weight, and Body Mass Index (BMI)

Even though BMI is an indicator of obesity that can increase the risk of many health problems, such as hypertension and diabetes, BMI can forecast the beginning of cardiovascular diseases [25,35].

Gender

Gender influences several vital signs in the human body [1], and researchers in [36] discovered no binding effect for changes in PPG signals due to gender. Regitz-Zagrosek et al. [37] said that cardiovascular events and heart failure are different for men and women, and Al-Fahoum et al. [12] found that men and women had different results during and before exercise.

Blood Pressure

Arza et al. [38] and Samria et al. [39] found a relationship between the PPG signal and blood pressure. Blood pressure correlates well with some PPG features.

Respiratory Problems

Respiration is a vital sign closely linked to the heart and lungs because of interactions between the two [11]. Monitoring respiratory activity may reveal important information that may aid in diagnosing CVD [3].

Smoking

Clair et al. statistically analyzed the relationship between smoking and CVDs risk. Their study revealed that quitting smoking can potentially decrease CVDs risk [40].

Physical Activity

Increasing physical activity improves cardiorespiratory fitness, which reflects favorably on the cardiovascular system by decreasing the risk of CVDs [41].

Other Variables (Diabetes, Kidney Failure, and Pregnancy)

Wannamethee et al. studied the effect of diabetes on the cardiovascular system and found that type I diabetes can influence the risk of coronary heart disease [42]. On the other hand, kidney failure and pregnancy were included in this study.

2.2.2. Pulse Oximetry

Arterial Oxygen Saturation (SpO₂)

SpO₂ is between 97% and 99% in healthy individuals and is clinically acceptable at 95%. If the SpO₂ value exceeds 95%, it may indicate an inadequate oxygen supply or hypoxia.

Pulse Rate

The average pulse rate is between 60 and 100 beats per minute. Caffeine, exercise, stress, certain medications, and other things can speed or slow the pulse rate.

2.2.3. PPG Signal

The PPG signal is susceptible to the subject's movements and breathing activity. These artifacts can change the baseline of the signal, causing shifts and offsets. In order to remove different artifacts and obtain a pure pulse wave, an elliptic band pass filter with a sampling frequency of 120 Hz, a lower cut-off frequency of 0.6 Hz, and an upper cut-off frequency of 15 Hz was used to eliminate the breathing effect.

The time-domain analysis of each pulse wave is the most important part of extracting features from the PPG signal. For this analysis, it is a must to find several time intervals and figure out where the peaks are in the signal [43]. The MATLAB software analyzed the pure PPG signal by taking out the following time-domain features:

Systolic Amplitude

Figure 2 shows the systolic amplitude that is the highest and the first peak of the pulse. Elgendi [44] mentioned that the systolic amplitude was found to be directly proportional to local vascular distensibility, while Awad et al. [10] stated that finger PPG systolic amplitude has a low correlation to the systemic vascular resistance (SVR).

Peak-to-Peak Interval (ΔT)

The time interval between two sequential systolic peaks can be seen in Figure 2. This interval is approximately the same as the R-R interval in ECG.

Pulse Interval

Elgendi defined the pulse interval as the time allocated between two consecutive minimums of the pulse [44]; see Figure 2. It relates to the time required for systole and diastole to be completed.

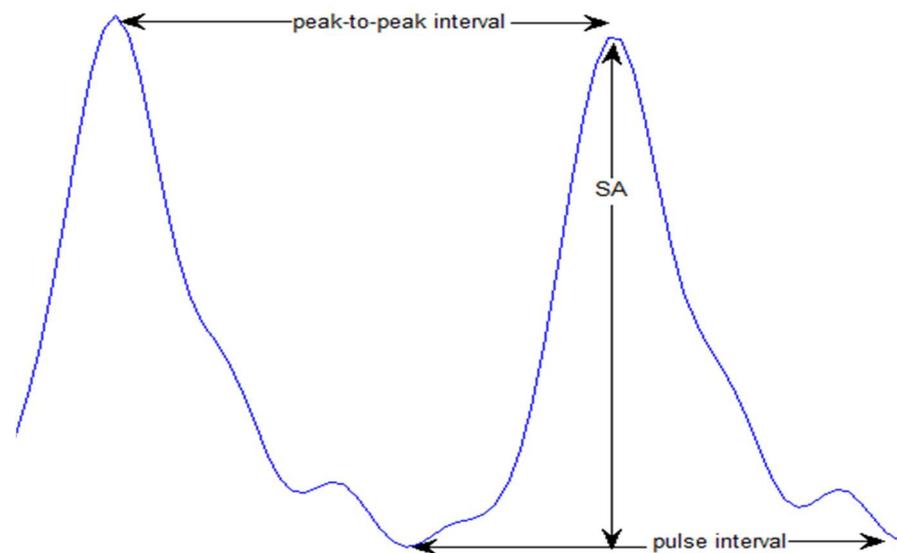


Figure 2. A display of SA along with peak-to-peak and pulse intervals.

Crest Time (CT)

Crest time is the interval between the start of the pulse wave and the time at the systolic peak. Angius et al. defined it as the amount of time needed for the fast ejection phase [9].

Pulse Width (PW)

As seen in Figure 3, it is the entire width of the pulse at the half height of the systolic peak. Awad et al. [10] calculated the SVR depending on the pulse width of the finger and PPG signals; they suggested that the higher the pulse width, the higher the SVR.

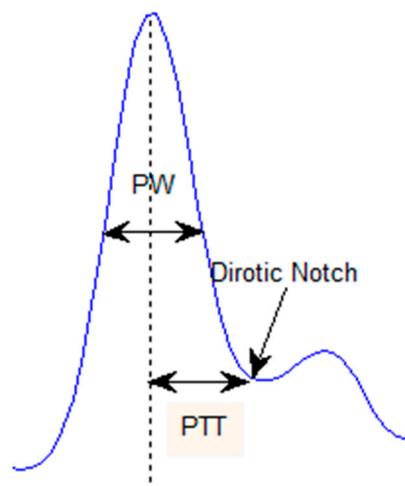


Figure 3. A display of PW, PTT and dicrotic notch.

Dicrotic Notch

Figure 3 points to the dicrotic notch point separating the systolic and diastolic phases. Angius et al. stated that sometimes the second or diastolic peak is absent in the pulse wave, especially in older adults and CVDs patients; in this case, the dicrotic notch substitutes for this peak [9].

Diastolic Time

It is the time interval from the dicrotic notch to the end of the pulse wave; it relates to the time required for the diastolic phase to be completed.

Pulse Transit Time (PTT)

Figure 3 shows the transit time of the pulse and the time between the systolic peak and the diastolic notch. According to Peulić et al. [11], age is negatively proportional to PTT.

Total Area and Inflection Point Area Ratio (IPA)

The total area is the whole area under the pulse wave curve that can be divided into two areas: the systolic and diastolic areas. These two areas are separated at the dicrotic notch. The total area parameter was suggested to indicate arterial stiffness [45]. The ratio of the diastolic area-to-systolic area is called IPA, and it can be used to measure total peripheral resistance [46].

Augmentation Indices (AI) and Time Ratios

Table 1 shows four different augmentation indices and three different time ratios that can be calculated from the PPG waveform. Augmentation indices may be used as a measure of systemic arterial stiffness [47,48]. On the other hand, Angius et al. [9] found a higher mean value of relative crest time for CVD patients.

Table 1. Augmentation Indices and Time Ratios.

AI/Time Ratio	Mathematical Formula
AI of Amplitude	(SA-Dicrotic Notch Amplitude)/SA
Reflection Index (RI)	Dicrotic Notch Amplitude/SA
AI of Time	PTT/Pulse Interval
Stiffness Index (SI)	Subject’s Height/PTT
Relative Crest Time (T1 Ratio)	CT/ ΔT
T2 Ratio	Systolic Time/ ΔT
T3 Ratio	PTT/ ΔT

APG Parameters and Their Ratios

In Figure 4, the parameters taken from the APG signal are shown. These parameters can be used to measure vascular aging. Baek et al. [49] studied some ratios between APG parameters and found them to be related to the cardiovascular system and age.

2.3. Feature Selection and Classification

Selecting features or attributes is searching for and selecting the best set of features that provides the highest classification accuracy. WEKA software allows for the selection of features from the dataset by applying two-step procedures: an attribute evaluator and a search method. In this study, the classifier subset evaluator method was used as the attribute evaluator. This method uses a classifier to estimate the best set of attributes. The classifier subset evaluator figures out how valuable a subset of attributes is by looking at how well each feature predicts on its own and how much it overlaps with other features. Subsets of attributes that are highly correlated with the class, while having low inter-correlation, are preferred. At the same time, the searching method was greedy stepwise, which performed a forward or backward search through the space of attribute subsets. For each classifier, all the features were analyzed, and the set of features that maximized the accuracy of the classifier was selected.

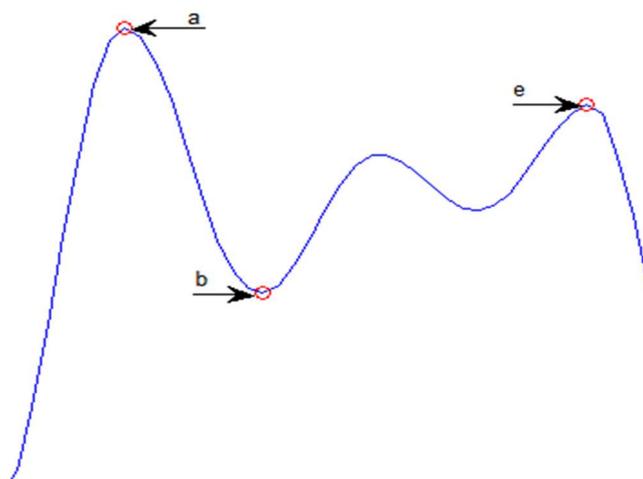


Figure 4. Parameters of the APG signals are: the “a” wave is the early systolic positive wave or the first positive peak of APG, the “b” wave is the early systolic negative wave or the first negative peak of APG, and the “e” wave characterizes a diastolic notch.

The frequencies of three CVD cases in the collected data, namely AF, DVT, and HF, were relatively small, compared to the other two cases, leading to an imbalanced class frequency problem. Under such situations, classifiers ignore the lower frequency classes and treat them as errors. In order to force the classifier to pay more attention to lower-frequency classes, we replicated the low-frequency classes.

2.4. Machine Learning Classification Algorithm

The proposed algorithm was conducted in four steps, as shown in Figure 5. In the first step, data for healthy and CVD patients were collected. Information about the person being studied’s age, gender, health, and PPG signal were among the things that were collected. MATLAB extracted time-domain features from each PPG signal. In the third step, a feature selection-based classifier was used. This was performed using the Waikato Environment for Knowledge Analysis (WEKA) software. This step involved two stages: the first distinguishes between healthy and CVD subjects and the second classifies five CVD cases. A feature selection process was used at both stages to pull out each classifier’s best, most accurate features. Finally, the classifier with the highest accuracy was selected. From another perspective, statistical analysis by MINITAB software was performed to compare the time-domain features of healthy and CVD subjects. The difference between the two samples was interpreted by applying a *t*-test, which tests the null hypothesis of whether the means of the two samples are equal.

In the first step of the classification process, each time-domain characteristic’s mean and standard deviation was utilized. In the second stage, when the search space became more extensive, the search strategy could not choose the best subset of features.

Due to the fact that AF, DVT, and HF occurred less frequently than the other two cardiac conditions, there was an imbalance in class frequency. In such cases, classifiers disregarded low-frequency classes as mistakes and discarded them. Classes with low frequency were replicated, so the classifier would pay greater attention to them. As indicated in the literature, each participant was assigned a unique PPG signal and treated as a new patient.

A *p*-value between 0 and 1 for the *t*-test on the mean of time-domain features for healthy and CVD patients demonstrates the result’s significance [8].

The significance level was set at 0.05; a *p*-value > 0.05 indicates no difference in the means of the two samples, and a *p*-value below 0.05 indicates a significant difference between the means of the two samples.

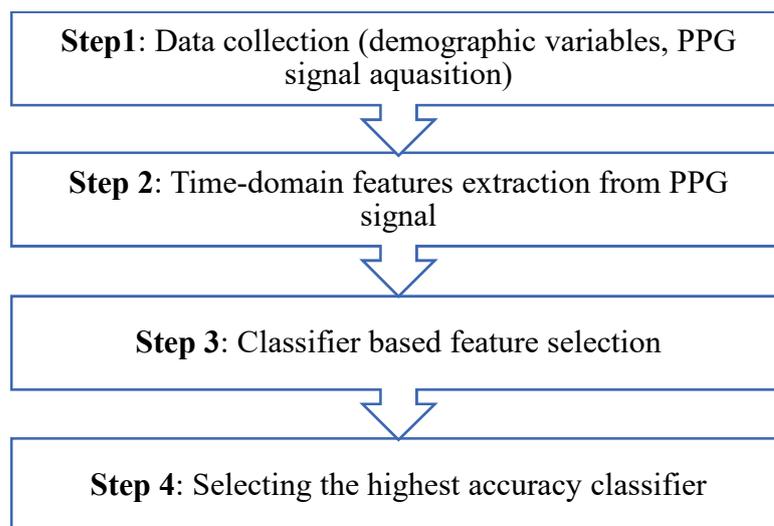


Figure 5. Schematic diagram of the proposed classification algorithm.

Accordingly, this study induced two stages of classification. The first one classifies data into two distinct categories: healthy and CVD. In the first step of classifying the subjects, each time-domain feature’s mean and standard deviation were used. The second stage classifies CVD cases into five classes, where only the feature mean value was used, since the search method could not select the best subset of features, due to the increased search space.

In this study, seven types of classifiers were considered, i.e., decision trees (J48 and random forest), rule-based (J-Rip and PART), artificial neural network/ANN (multilayer perceptron), K-nearest neighbor (KNN), and Bayesian (Naïve Bayes) classifiers, in both stages. The attribute selection tool in WEKA selects a set of features for each classifier. This set of features is considered the best subset that improves the classifier’s performance.

3. Results

Each classifier’s subset of features was used to sort features into groups, and the most accurate classifier was chosen. Additionally, 10-fold cross-validation was used in this step. Tables 2 and 3 show each classifier’s set of features and their accuracies in the first stage. Tables 4 and 5 depict each classifier’s features and their accuracies in the second stage of the classification process. In both stages: naïve Bayes gave the highest accuracy in both stages (94.44% and 89.37%, respectively). Naïve Bayes also provided accuracies of 66.12% and 62.28%, based on PPG features only.

Table 2. Selected features by each classifier in the first stage.

Classifier	Selected Demographic Features	Selected PPG Features
J48	Age, weight	Pulse rate, PW, diastolic time, a, variance of e
Random forest		Pulse rate, variance of ΔT
J-Rip	Age	Pulse rate, variance of pulse interval, notch, variance of e
PART	Age, weight	Pulse rate, variance of total area, diastolic time, a, variance of e
Naïve Bayes	Age, smoking, respiratory problem, others	SA, variance of diastolic time
ANN	Age, height, BMI, gender, smoking, respiratory problem, blood pressure, others	SA, variance of SA, Notch, variance of diastolic time, AI of amplitude
KNN	Age	Variance of total area

Table 3. The accuracy of seven classifiers in the first classification stage.

Classifier	Accuracy (%)
J48	89.10
Random forest	62.94
J-Rip	85.15
PART	85.33
Naïve Bayes	94.44
ANN	89.56
KNN	82.08

Table 4. Selected features by each classifier in the second stage.

Classifier	Selected Demographic Features	Selected PPG Features
J48	Age, weight, BMI	Pulse rate, SpO ₂ , SA, diastolic time, T1, b, e, e/a
Random forest		ΔT
J-Rip	Age, weight, height, BMI, smoking	SpO ₂ , Notch, SI, b/a
PART	Age, weight, BMI, smoking	Total area, A1, T2, AI of amplitude, b, e, e/a
Naïve Bayes	Age, height, gender, smoking, respiratory problem, blood pressure	SpO ₂ , ΔT, Notch, total area, b/a, e, e/a, SI
ANN	Age, height, gender, physical activity, smoking, respiratory problem, blood pressure	CT, Notch, diastolic time, SI, a, b, b/a, e, e/a, (b-e)/a
KNN	Age, height	Notch

Table 5. The accuracy of seven classifiers in the second classification stage.

Classifier	Accuracy (%)
J48	59.36
Random forest	34.12
J-Rip	69.07
PART	72.07
Naïve Bayes	89.37
ANN	88.52
KNN	83.64

Sometimes it is not favorable to depend only on accuracy to evaluate the classifier. The recall is another evaluator of the classifier. The better the classifier, the higher the recall. The recall values of naïve Bayes in both stages were 0.882 in the first one and 0.806 in the second one.

4. Discussion

Regarding the best subset of features selected by each classifier, age dominates each classifier’s selection. This result is the same as other studies, which said that age was an essential factor in how stiff the arteries were [10]. Likewise, some CVD risk factors, such as smoking, respiratory problems, hypertension, and diabetes, appeared repeatedly in the selected feature set.

Variability analysis of PPG time-domain features assesses their degree of variation over time. These features’ variance and their means helped differentiate healthy PPG signals from CVD signals. Naïve Bayes selected the mean systolic amplitude. The fact that arterial stiffness in CVD cases changes the amplitudes of the PPG signal matches this selection. Additionally, the classifier selected the variance of diastolic time within the signal. Based on the 61% accuracy, these two features can detect CVD occurrence.

Since the variance of PPG features over time among CVD cases did not improve the classifier’s performance, only the mean of each feature represented the PPG signal. The

inability to deal with the variance may be related to the fact that the search method could not select the best subset of features due to the increased search space. Arterial stiffness changes both the amplitudes and time intervals of the PPG signal, and each type of CVD affects the PPG amplitudes and intervals. So, in the second classification stage, naïve Bayes selected the CT, Notch, diastolic time, and SI, along with the APG parameters and ratios (a , b , b/a , e , e/a , $(b-e)/a$).

In this study, the differences between CVDs affect the CT, notch amplitude, and diastolic time selection. Finally, the SI and APG parameters are correlated with arterial stiffness, so selecting these features may relate to CVDs that include a blockage of vessels, such as CVA, DVT, and ACS.

From another point of view, naïve Bayes provides accuracies of 66.12% and 62.28%, based on the PPG features alone. The results indicate a significant effect of PPG features in classification. Both accuracy and recall in the two stages indicate an excellent evaluation of the constructed naïve Bayes classifier. The quality of the classification models can be measured by several performance measures with frequently contradictory outcomes. In [46], they compared multiple levels using various performance measures and machine learning categorization techniques. In each instance, well-established and defined techniques were utilized for the machine-learning tasks. Three datasets (acute and aquatic toxicities) were compared, and the robust yet sensitive sum of ranking differences (SRD) and analysis of variance (ANOVA) were used to evaluate the data. The effects of dataset composition (balanced versus unbalanced) and two-class versus multiclass classification scenarios were also investigated. Most performance indicators are sensitive to dataset composition, particularly in two-class classification issues. The ideal machine-learning algorithm is also very dependent on the dataset's makeup [4]. In [50–53], the outcomes of regression performance metrics are discussed in greater detail. These studies show how important it is to validate machine learning models, which tend to work better than “traditional” regression techniques. Based on the difference between the training set and the test set's coefficients of determination, the most reliable models are principal components and partial least squares regression.

5. Conclusions

As datasets, the mean and variance of the 24-time-domain feature extracted from PPG signals, demographic factors, and some information about health status were used to identify the difference between healthy human subjects and CVD patients. To classify the five types of CVDs, on the other hand, only the mean of the characteristics derived from the PPG signal and other covariates was used. In all instances, seven distinct classifiers [15,26–33] were utilized to apply feature selection and classification techniques. In the two classification stages, the naïve Bayes classifier scored the highest accuracies at 94.44% and 89.37%, respectively.

This study shows that the PPG signal can be used to find CVD by using a minimum number of PPG time-domain features, demographic factors, and some health status information. Based on these results, future work will use machine learning and deep learning algorithms to create other classification methods.

Author Contributions: Conceptualization, A.S.A.F. and A.O.A.A.-H.; methodology, A.S.A.F.; software, A.O.A.A.-H.; validation, A.S.A.F. and H.A.A.; formal analysis, A.S.A.F.; investigation, A.S.A.F.; resources, A.S.A.F.; data curation, A.O.A.A.-H.; writing—original draft preparation, A.S.A.F., A.O.A.A.-H. and H.A.A.; writing—review and editing, A.S.A.F.; visualization, A.O.A.A.-H.; supervision, A.O.A.A.-H.; project administration, A.S.A.F.; funding acquisition, A.S.A.F. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would like to thank the deanship of research and higher studies at Yarmouk University for its support and funding.

Institutional Review Board Statement: The study was approved by the institutional review board at Jordan University of Science and Technology IRB-2019-07.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is unavailable due to privacy of the hospital requirements and ethical restrictions.

Conflicts of Interest: The authors declare that they have no conflict of interest to report regarding the present study.

References

1. Labarthe, D. *Epidemiology and Prevention of Cardiovascular Diseases*, 2nd ed.; Jones and Bartlett Publishers: Burlington, MA, USA, 2010.
2. Jordanian Ministry of Health. Periodic-Newsletters, Amman. 2020. Available online: <http://www.moh.gov.jo/EN/Pages/Periodic-Newsletters.aspx> (accessed on 18 September 2020).
3. Jain, P.; Tiwari, A. Heart monitoring systems: A review. *Comput. Biol. Med.* **2014**, *54*, 1–13. [[CrossRef](#)]
4. Rácz, A.; Bajusz, D.; Héberger, K. Multi-Level Comparison of Machine Learning Classifiers and Their Performance Metrics. *Molecules* **2019**, *24*, 2811. [[CrossRef](#)]
5. Al-Fahoum, A.; Khadra, L. Combined Bispectral and Bicoherency approach for Catastrophic Arrhythmia Classification. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2005; pp. 332–336. [[CrossRef](#)]
6. Al-Fahoum, A.S.; Qasaimeh, A.M. A practical reconstructed phase space approach for ECG arrhythmias classification. *J. Med. Eng. Technol.* **2013**, *37*, 401–408. [[CrossRef](#)]
7. Al-Fahoum, A.; Al-Fraihat, A.; Al-Araida, A. Detection of cardiac ischaemia using bispectral analysis approach. *J. Med. Eng. Technol.* **2014**, *38*, 311–316. [[CrossRef](#)] [[PubMed](#)]
8. Héberger, K. Sum of ranking differences compares methods or models fairly. *TrAC Trends Anal. Chem.* **2010**, *29*, 101–109. [[CrossRef](#)]
9. Angius, G.; Barcellona, D.; Cauli, E.; Meloni, L.; Raffo, L. Infarction and antiphospholipid syndrome: A first study on finger PPG waveforms effects. *Comput. Cardiol.* **2012**, *12*, 517–520.
10. Awad, A.; Haddadin, A.; Tantawy, H.; Badr, T.; Stout, R.; Silverman, D.; Shelley, K. The relationship between the photoplethysmographic waveform and systemic vascular resistance. *J. Clin. Monit. Comput.* **2007**, *21*, 365–372. [[CrossRef](#)]
11. Peulic, A.; Milojevic, N.; Jovanov, E.; Radovic, M.; Saveljic, I.; Zdravkovic, N.; Filipovic, N. Modeling of arterial stiffness using variations of pulse transit time. *Comput. Sci. Inf. Syst.* **2013**, *10*, 547–565. [[CrossRef](#)]
12. Al Fahoum, A.S.; Al Zaben, A.; Seafan, W. A multiple signal classification approach for photoplethysmography signals in healthy and athletic subjects. *Int. J. Biomed. Eng. Technol.* **2015**, *17*, 1–23. [[CrossRef](#)]
13. Aboy, M.; McNames, J.; Thong, T.; Tsunami, D.; Ellenby, M.; Goldstein, B. An Automatic Beat Detection Algorithm for Pressure Signals. *IEEE Trans. Biomed. Eng.* **2005**, *52*, 1662–1670. [[CrossRef](#)] [[PubMed](#)]
14. Solosenko, A.; Petrenas, A.; Marozas, V. Photoplethysmography-Based Method for Automatic Detection of Premature Ventricular Contractions. *IEEE Trans. Biomed. Circuits Syst.* **2015**, *9*, 662–669. [[CrossRef](#)]
15. Yousefi, M.R.; Khezri, M.; Bagheri, R.; Jafari, R. Automatic detection of premature ventricular contraction based on photoplethysmography using chaotic features and high order statistics. In Proceedings of the IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rome, Italy, 11–13 June 2018.
16. Polania, L.F.; Mestha, L.K.; Huang, D.T.; Couderc, J.P. Method for classifying cardiac arrhythmias using photoplethysmography. In Proceedings of the International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), Milano, Italy, 25–29 August 2015.
17. Al-haija, A.A.; Alshraideh, H.; Al-Fahoum, A. Cardiorespiratory Disorders Detection Using Photoplethysmography. In Proceedings of the 81st IASTEM International Conference, Phuket, Thailand, 9–10 October 2017; pp. 65–69.
18. Nano, M.; Papini, G.; Fonseca, P. Comparing inter beat and inter pulse intervals from ECG and PPG signals. In Proceedings of the Biomedica Summit, Eindhoven, The Netherlands, 9–10 May 2017.
19. Papini, G.B.; Fonseca, P.; Aubert, X.L.; Overeem, S.; Bergmans, J.W.; Vullings, R. Photoplethysmography beat detection and pulse morphology quality assessment for signal reliability estimation. In Proceedings of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju Island, Republic of Korea, 11–15 July 2017. [[CrossRef](#)]
20. Prabhakar, S.K.; Rajaguru, H.; Lee, S.-W. Metaheuristic-Based Dimensionality Reduction and Classification Analysis of PPG Signals for Interpreting Cardiovascular Disease. *IEEE Access* **2019**, *7*, 165181–165206. [[CrossRef](#)]
21. Fayyad, U.; Uthurusamy, R. Data mining and knowledge discovery in databases. *Commun. ACM* **1996**, *39*, 24–26. [[CrossRef](#)]
22. Ahmed, K.; Abdullah-Al-Emran, A.-A.; Jesmin, T.; Mukti, R.F.; Rahman, Z.; Ahmed, F. Early Detection of Lung Cancer Risk Using Data Mining. *Asian Pac. J. Cancer Prev.* **2013**, *14*, 595–598. [[CrossRef](#)]
23. Pilt, K.; Ferenets, R.; Meigas, K.; Lindberg, L.-G.; Temitski, K.; Viigimaa, M. New Photoplethysmographic Signal Analysis Algorithm for Arterial Stiffness Estimation. *Sci. World J.* **2013**, *2013*, 169035. [[CrossRef](#)] [[PubMed](#)]
24. Yousef, Q.; Reaz, M.B.I.; Ali, M.A.M. The Analysis of PPG Morphology: Investigating the Effects of Aging on Arterial Compliance. *Meas. Sci. Rev.* **2012**, *12*, 266–271. [[CrossRef](#)]

25. Tybor, D.J.; Lichtenstein, A.H.; Dallal, G.E.; Daniels, S.R.; Must, A. Independent effects of age-related changes in waist circumference and BMI z scores in predicting cardiovascular disease risk factors in a prospective cohort of adolescent females. *Am. J. Clin. Nutr.* **2010**, *93*, 392–401. [[CrossRef](#)] [[PubMed](#)]
26. Allen, J.; Liu, H.; Iqbal, S.; Zheng, D.; Stansby, G. Deep learning-based photoplethysmography classification for peripheral arterial disease detection: A proof-of-concept study. *Physiol. Meas.* **2021**, *42*, 054002. [[CrossRef](#)] [[PubMed](#)]
27. Shabaan, M.; Arshid, K.; Yaqub, M.; Jinchao, F.; Zia, M.S.; Bojja, G.R.; Iftikhar, M.; Ghani, U.; Ambati, L.S.; Munir, R. Survey: Smartphone-based assessment of cardiovascular diseases using ECG and PPG analysis. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 1–16. [[CrossRef](#)]
28. Olsen, M.; Zeitzer, J.M.; Richardson, R.N.; Davidenko, P.; Jennum, P.J.; Sorensen, H.B.D.; Mignot, E. A Flexible Deep Learning Architecture for Temporal Sleep Stage Classification Using Accelerometry and Photoplethysmography. *IEEE Trans. Biomed. Eng.* **2022**, *70*, 228–237. [[CrossRef](#)]
29. Schruppf, F.; Serdack, P.R.; Fuchs, M. Regression or Classification? In Reflection on BP prediction from PPG data using Deep Neural Networks in the scope of practical applications. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LO, USA, 18–24 June 2022; pp. 2171–2180. [[CrossRef](#)]
30. Fakhry, M.; Brery, A.F. A Comparison Study on Training Optimization Algorithms in the biLSTM Neural Network for Classification of PCG Signals. In Proceedings of the 2022 2nd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco, 3–4 March 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6. [[CrossRef](#)]
31. Bhadri, K.; Karnik, N.; Dhattrak, P. Current Advancements in Cardiovascular Disease Management using Artificial Intelligence and Machine Learning Models: Current Scenario and Challenges. In Proceedings of the 2022 10th International Conference on Emerging Trends in Engineering and Technology-Signal and Information Processing (ICETET-SIP-22), Nagpur, India, 29–30 April 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6. [[CrossRef](#)]
32. Neha; Sardana, H.K.; Kanwade, R.; Tewary, S. Arrhythmia detection and classification using ECG and PPG techniques: A review. *Phys. Eng. Sci. Med.* **2021**, *44*, 1027–1048. [[CrossRef](#)] [[PubMed](#)]
33. Khan, M.U.; Aziz, S.; Iqtidar, K.; Zaher, G.F.; Alghamdi, S.; Gull, M. A two-stage classification model integrating feature fusion for coronary artery disease detection and classification. *Multimed. Tools Appl.* **2022**, *81*, 13661–13690. [[CrossRef](#)]
34. Beurer PO 80 Pulse Oximeter. Available online: <https://www.beurer.com/web/gb/products/medical/pulse-oximeter/po-80.php> (accessed on 8 January 2023).
35. Freedman, D.; Ogden, C.; Kit, B. Interrelationships between BMI, skinfold thicknesses, percent body fat and cardiovascular disease risk factors among U.S. children and adolescents. *BMC Pediatr.* **2015**, *15*, 188. [[CrossRef](#)] [[PubMed](#)]
36. Nilsson, L.; Goscinski, T.; Johansson, A.; Lindberg, L.-G.; Kalman, S. Age and Gender Do Not Influence the Ability to Detect Respiration by Photoplethysmography. *J. Clin. Monit. Comput.* **2006**, *20*, 431–436. [[CrossRef](#)] [[PubMed](#)]
37. Regitz-Zagrosek, V.; Lehmkuhl, E.; Weickert, M. Gender differences in the metabolic syndrome and their role for cardi-ovascular disease. *Clin. Res. Cardiol.* **2006**, *95*, 136–147. [[CrossRef](#)] [[PubMed](#)]
38. Arza, A.; Lazaro, J.; Gil, E.; Laguna, P.; Aguilo, J.; Bailon, R. Pulse Transit Time and Pulse Width as Potential Measure for Estimating Beat-to-Beat Systolic and Diastolic Blood Pressure. *Comput. Cardiol.* **2013**, *40*, 887–890.
39. Samria, R.; Jain, R.; Jha, A.; Saini, S.; Chowdhury, S. Noninvasive cuffless estimation of blood pressure using photoplethysmography without electrocardiograph Measurement. In Proceedings of the International Conference IEEE Region 10 Symposium, Kuala Lumpur, Malaysia, 14–16 April 2014.
40. Clair, C.; Rigotti, N.A.; Porneala, B.; Fox, C.S.; D’Agostino, R.B.; Pencina, M.J.; Meigs, J.B. Association of Smoking Cessation and Weight Change with Cardiovascular Disease Among Adults with and Without Diabetes. *JAMA* **2013**, *309*, 1014–1021. [[CrossRef](#)]
41. Durand, C.P.; Andalib, M.; Dunton, G.F.; Wolch, J.; Pentz, M.A. A systematic review of built environment factors related to physical activity and obesity risk: Implications for smart growth urban planning. *Obes. Rev.* **2011**, *12*, e173–e182. [[CrossRef](#)] [[PubMed](#)]
42. Wannamethee, S.G.; Shaper, A.G.; Whincup, P.; Lennon, L.; Sattar, N. Impact of Diabetes on Cardiovascular Disease Risk and All-Cause Mortality in Older Men. *Arch. Intern. Med.* **2011**, *171*, 404–410. [[CrossRef](#)]
43. Kuntamalla, S.; Reddy, L.R.G. An Efficient and Automatic Systolic Peak Detection Algorithm for Photoplethysmographic Signals. *Int. J. Comput. Appl.* **2014**, *97*, 18–23. [[CrossRef](#)]
44. Elgendi, M. On the Analysis of Fingertip Photoplethysmogram Signals. *Curr. Cardiol. Rev.* **2012**, *8*, 14–25. [[CrossRef](#)]
45. Usman, S.; Reaz, M.; Ali, M. Repeated measurement analysis of the area under the curve of photoplethysmogram among diabetic patients. *Life Sci. J.* **2011**, *10*, 532–539.
46. Wang, J. *Data Mining*; Idea Group Pub: Hershey, PA, USA, 2003.
47. Manimegalai, P.; Jacob, D.; Thanushkodi, K. An Early Prediction of Cardiac Risk using Augmentation Index Developed based on a Comparative Study. *Int. J. Comput. Appl.* **2012**, *49*, 27–32.
48. Jayasree, V.; Sandhya, T.; Radhakrishnan, P. Non-invasive Studies on Age Related Parameters Using a Blood Volume Pulse Sensor. *Meas. Sci. Rev.* **2008**, *8*, 82–86. [[CrossRef](#)]
49. Beak, H.; Kim, J.; Kim, Y.; Lee, H.; Park, K. Second derivative of photoplethysmography for estimating vascular aging. In Proceedings of the 6th International Special Topic Conference on Information Technology Applications in Biomedicine, Tokyo, Japan, 8–11 November 2007.
50. Kruschke, J. Bayesian estimation supersedes the t test. *J. Exp. Psychol. Gen.* **2013**, *142*, 573–603. [[CrossRef](#)]

51. Rácz, A.; Bajusz, D.; Héberger, K. Consistency of QSAR models: Correct split of training and test sets, ranking of models and performance parameters. *SAR QSAR Environ. Res.* **2015**, *26*, 683–700. [[CrossRef](#)]
52. Héberger, K.; Rácz, A.; Bajusz, D. Which Performance Parameters Are Best Suited to Assess the Predictive Ability of Models? In *Advances in QSAR Modeling*; Roy, K., Ed.; Springer: Cham, Switzerland, 2017; pp. 89–104. [[CrossRef](#)]
53. Rácz, A.; Bajusz, D.; Héberger, K. Modelling methods and cross-validation variants in QSAR: A multi-level analysis. *SAR QSAR Environ. Res.* **2018**, *29*, 661–674. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.