Hypertension Assessment Using Photoplethysmography: A Risk Stratification Approach

Yongbo Liang 1,2, Zhencheng Chen 1,*, Rabab Ward 2 and Mohamed Elgendi 2,3,4,*

1 School of Electronic Engineering and Automation, Guilin University of Electronic Technology, Guilin 541004, China; liangyongbo001@gmail.com
2 School of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC V6T 1Z4, Canada; rababw@ece.ubc.ca
3 Faculty of Medicine, University of British Columbia, Vancouver, BC V1Y 1T3, Canada
4 BC Children’s and Women’s Hospital, Vancouver, BC V6H 3N1, Canada

* Correspondence: chenzhcheng@163.com (Z.C.); melgendi@bcchr.ubc.ca (M.E.); Tel.: +1-604-600-4139 (M.E.)

Received: 14 November 2018; Accepted: 14 December 2018; Published: 21 December 2018

Abstract: Hypertension is a common chronic cardiovascular disease (CVD). Early screening and diagnosis of hypertension plays a major role in its prevention and in the control of CVDs. Our study discusses the early screening of hypertension while using the morphological features of photoplethysmography (PPG). Numerous morphological features of PPG and its derivative waves were defined and extracted. Six types of feature selection methods were chosen to screen and evaluate these PPG morphological features. The optimal features were comprehensively analyzed in relation to the physiological processes of the cardiovascular circulatory system. Particularly, the intrinsic relation and physiological significance between the formation process of systolic blood pressure (SBP) and PPG morphology features were analyzed in depth. A variety of linear and nonlinear classification models were established for the comparison trials. The F1 scores for the normotension versus prehypertension, normotension and prehypertension versus hypertension, and normotension versus hypertension trials were 72.97%, 81.82%, and 92.31%, respectively. In summary, this study established a PPG characteristic analysis model and established the intrinsic relationship between SBP and PPG characteristics. Finally, the risk stratification of hypertension at different stages was examined and compared based on the optimal feature subset.

Keywords: Hypertension; systolic blood pressure; photoplethysmograph; feature selection; risk classification

1. Introduction

In recent years, instances of cardiovascular diseases (CVDs) have been increasing worldwide due to a lack of awareness and a low rate of diagnosis and treatment. Meanwhile, the mortality rate of CVDs has overtaken that of cancer, making CVDs the leading cause of death in humans and, accordingly, a major public health problem [1]. The unintentional and unpredictable occurrence of acute CVDs [2], such as coronary artery disease, stroke, heart failure, and myocardial infarction, can have irreversible consequences and may produce many sequelae; whereas, chronic CVDs, such as hypertension, hyperlipidemia, and atherosclerosis, cause slow deterioration, thereby leading to a greater potential for delayed treatment as well as the need for lifelong medication. Many CVDs require long-term health care and protection, which consumes large amounts of medical and health resources; consequently, such CVDs represent a global health challenge.
As an important physiological index of the cardiovascular system, blood pressure (BP) plays a major role in predicting the occurrence of acute and chronic CVDs [3,4]. Therefore, the early screening and evaluation of patients with hypertension or prehypertension is of great practical significance [5,6]. Hypertension [7] in adults over 18 years of age is defined as systolic or diastolic BP that is consistently higher than an acceptable normal value (currently 139 mmHg systolic, 89 mmHg diastolic). Different high BP levels are accompanied by different risks; therefore, hypertension has been subdivided into stage 1 hypertension and stage 2 hypertension. The criterion for stage 1 hypertension is systolic blood pressure (SBP) that is in the range of 140 to 159 mmHg or diastolic blood pressure (DBP) in the range of 90 to 99 mmHg. The criterion for stage 2 hypertension is SBP exceeding 160 mmHg or DBP exceeding 100 mmHg. The measures used to determine normotension are based on the evolutionary trend of human BP and have been subdivided into normotension and prehypertension. The criterion for normotension is SBP ranging between 90 to 119 mmHg and DBP ranging between 60 to 79 mmHg concurrently. The criterion for prehypertension is SBP ranging from 120 to 139 mmHg or DBP ranging from 80 to 89 mmHg. It is well known that BP fluctuates with age. For those with hypertension, without proper control and management, they will most likely need treatment. Therefore, the management and control of hypertension is critical to reducing its severity; additionally, numerous non-drug methods, such as lifestyle changes and exercise, can significantly improve BP, returning it to normal levels.

BP is currently measured with an electronic sphygmomanometer [8]. This measurement method has been widely recognized and popularized over the last century of development and it has played a major role in the control of CVDs. However, both the Korotkoff sound [9] method and the oscillometric method [10] require cuffs and pressure to the forearm when BP is measured, which means that the measurement can easily be affected by operation and use conditions, such as the operation of the cuff, sitting posture, and exercise. In addition, it also has a definite white coat phenomenon (BP is higher when it is taken in a medical setting than it is when taken at home) for some patients [3]. Therefore, new cuff-less hypertension screening and BP detection technologies are under development [11].

Photoplethysmography (PPG) provides abundant physiological information [12] about the operation of the cardiovascular circulation system at a low cost and with convenient signal acquisition; therefore, it is of great interest to researchers. A series of in-depth investigations [13–17] have been carried out to non-invasively assess BP, blood sugar, stiffness, and fatigue levels using PPG signals. The common PPG acquisition methods can be divided into transmission and reflection methods [18]. A PPG signal is extracted by an LED transmitter section, which generates a red or infrared light to illuminate the skin of the fingertip, earlobe, or forehead, and a photosensitive diode, which measures the time-varying light absorption of the tissue and it can reflect changes in blood volume. Therefore, PPG is the external manifestation and aggregated expression of various physiological processes in the cardiovascular circulation system. PPG is a high fusion signal that covers the activity of the heart’s systolic and diastolic periods [19] and the hemodynamic [20], hemorheological [21], and network information of the peripheral microcirculation system [22].

In recent years, more features of the PPG signal have been proposed and validated to predict BP or evaluate hypertension [23], such as pulse transit time [16], pulse wave velocity [23], augmentation index [5], stiffness index, PPG intensity ratio [24], and time-domain characteristics [25]. Meanwhile, arterial wave propagation theory [16] and PPG morphological theory have also been researched: The former is based on ECGs or other physiological signals and the PPG signal, while the latter is based on PPG morphological features. However, more research focused on the healthy subjects and the limited number of hypertensive subjects. Therefore, BP readings and PPG signals were collected for this study from relatively large number of healthy and hypertensive subjects [26,27].

To clarify hypertension risk stratification using only one non-invasive biosignal, the authors have conducted extensive research in this area [28–31]. We have found that two approaches can be implemented: a feature-based method [28,29] or a non-feature-based method [31]. For the sake of providing a simple model suitable for wearable devices, we have only focused on the feature-based
method. Therefore, our research design for this study can be outlined, as follows: (1) the physiological relevance of the PPG signal was comprehensively analyzed using the feature-based method (depending on the PPG waveform morphology), (2) the PPG features were selected and ranked using machine learning algorithms, (3) multiple linear and nonlinear models were investigated for hypertension risk stratification, and (4) a portable health device and matching app were developed based on the internet of things, wearable technologies, and Android platforms for ubiquitous hypertension screening and tracking.

2. Materials and Methods

2.1. Data Acquisition

In this study, a portable hardware platform was designed, consisting of a PPG probe and matching mobile phone app; the data were transmitted via Bluetooth. For the PPG probe, an MSP430FG4618 MCU was embedded on the board to configure the ADC and DAC and to fetch the data for transmission to the matching app via Bluetooth. The mobile phone app has several functional modules, namely the photoelectric control module, the waveform real-time display module, the sampling precision configuration module, and the data storage module. For the purposes of this study, the PPG signal acquisition was configured as an infrared (905-nm wavelength) transmission mode with a sampling rate of 1 kHz and a 12-bit ADC, and the hardware filter design was 0.5–12 Hz bandpass. Figure 1 shows a diagram of the data collection process and the customized hardware platform.

![Figure 1. The block diagram of data collection. (a) Data collection protocol; (b) Blood pressure (BP) collection; (c) photoplethysmography (PPG) collection and the customized hardware platform.](image-url)

Data collection was performed on 219 subjects recruited from Guilin People’s Hospital [26] and was conducted according to the ethics rules and regulations of both the hospital and the Guilin University of Electronic Technology in China after approval by the ethics committee. Participants who suffered from diseases, such as neurological disorders were excluded, as were those suffering from diabetes (38 subjects), cerebral infarction (20 subjects), cerebrovascular disease (37 subjects), or those who were taking hypotensive drugs during the experimental period. In the experiment,
participants first waited in a rest area for more than 30 min before entering the data collection room. After entering the room, they were seated in a chair with a backrest in the most comfortable posture and instructed to place their arms flat on a blank desktop. The operators simultaneously used a customized probe to collect the PPG signal from the participants’ left index fingers and an Omron 7201 electronic sphygmomanometer to determine BP from their right forearms. The PPG and BP acquisition process was completed within 3 min. Three 2.1-s-long PPG waveform data records were saved in the matching app for each participant. Additionally, physiological information from the participants, including gender, age, height, and weight, was also collected through the app, and disease records were obtained through the hospital’s electronic medical records. More details about the background and introduction of the project can be found in [26]. The dataset has been fully uploaded to the network and users can download it via this link [26].

Before the experiment, all of the participants were fully informed about the experimental content and their informed consent was obtained. In addition, before the data were archived, we removed personal identifiers. Ultimately, we generated a BP database that includes data from 124 subjects (age: 55 ± 16; Height: 163 ± 8 cm; Weight: 62 ± 15 kg; 65 males and 59 females). The risk stratification approach was implemented by dividing the dataset according to the hypertension classification criteria: hypertensive group (SBP ≥ 140 or DBP ≥ 90), prehypertensive group (SBP 120–140 mmHg or DBP 80–90 mmHg), and normotensive group (SBP < 120 and DBP < 80). The hypertensive group included 35 subjects; the prehypertensive group included 41 subjects; and, the normotensive group included 48 subjects.

2.2. PPG Signal Pre-Processing

The raw PPG signal was first pre-processed by a PPG signal quality evaluation [32], a bandpass filter, and a mathematical derivative [13]. The PPG signal quality evaluation was conducted via the skewness method for the three waveform records of each subject. Elgendi [32] compared eight different signal quality indices: $P_{SQI}$, skewness ($S_{SQI}$), kurtosis ($K_{SQI}$), entropy ($E_{SQI}$), signal-to-noise ratio ($N_{SQI}$), zero-crossing ($Z_{SQI}$), matching of multiple systolic wave detection algorithms ($M_{SQI}$), and relative power ($R_{SQI}$). For lengths of PPG waveforms between 2 s and 30 s, the $S_{SQI}$ method demonstrated better performance than the other methods ($P_{SQI}$, $K_{SQI}$, $E_{SQI}$, $N_{SQI}$, $M_{SQI}$, $Z_{SQI}$, and $R_{SQI}$). Using this method, the classifications of excellent or unfit waveforms were best when 3 s was used as the window of the PPG waveform segment. In this study, the $S_{SQI}$ method was used to evaluate the records, and only the segment with the highest $S_{SQI}$ value for each subject, which included the PPG feature analysis and the BP classification model, was selected. For all the selected waveform records, a fourth order 0.5 Hz–10 Hz bandpass Chebyshev II filter was used to filter noise and improve the PPG signal quality while, at the same time, the PPG features were kept undistorted [27]. Subsequently, two times the mathematical derivative of the PPG signal was implemented to acquire the velocity plethysmogram (VPG), the acceleration plethysmogram (APG), and the third derivative waveform (third derivative), all of which were used to define and extract the characteristics of the PPG signal.

2.3. Feature Extraction

A PPG signal contains information on the heart’s systolic and diastolic activity, hemodynamics, and hemorheology. Therefore, in a PPG signal and its derivatives, many potential physiological characteristics [12,33] can be defined or quantified. Some of these characteristics have been found and confirmed by a large number of researchers; for example, the characteristics of the systolic and diastolic periods, crest time, augmentation index, and large artery stiffness index [33]. There are other characteristics that have potential physiological feasibility, but these have not yet been confirmed in clinical data and have therefore not yet been valued.

In this study, a comprehensive investigation of the characteristics of the PPG signal and its derivatives was conducted according to the recommendation in [30]. For consistency and clarity, we have suggested the following waveform names: “Onset (O)”, “Systolic Peak (S)”, “Notch (N)”, and
“Diastolic Peak (D)” in the PPG waveform; “w wave (w), “x wave (x), “y wave (y), and “z wave (z)” in the VPG waveform; and “a wave (a), “b wave (b), “c wave (c), “d wave (d), and “e wave (e)” in the APG waveform. The b, c, and d waves appearing in the PPG and VPG waveforms were named “a_{-2}, “b_{-2}, “c_{-2}, “d_{-2}, and “e_{-2}, and “c_{-1}” and “d_{-1}”, respectively. All of the wave names are marked in Figure 2. The characteristics’ features based on their wave names were defined and include several types, such as time span (23 features), PPG amplitude (14 features), features of VPG and APG (10 features), waveform area (4 features), power area (15 features), ratio (43 features), and slope (16 features). A total of 125 features are defined and described in this paper.

2.3.1. Time Span

The time span features are expressed as their letters with a dash over each. For example, the $SD$ feature represents the time span from the systolic peak S to the diastolic peak D. A total of 23 time-span features were defined.

2.3.2. Features of PPG Amplitude

The S, N, D, $w_{-1}, a_{-2}, b_{-2}, and c_{-2}$ features were defined in the PPG waveform. These features represent the amplitude of the corresponding waveform from the PPG baseline. For example, the S amplitude feature represents the S height from the PPG baseline to the systolic peak S, and the N amplitude feature represents the N height from the PPG baseline to the diastolic notch N. Other features that are based on these features were also defined, such as N/S, $w_{-1}/S, b_{-2}/S$, and $c_{-2}/S$. A total of 14 PPG amplitude features were defined.

2.3.3. Features of VPG and APG

The $a, b, c, d, and e$ features were defined in the APG waveform (the automatic detection of these waves is discussed here: [34–37]), and the $w, x, y, and z$ features were defined in the VPG waveform. These features represent the amplitude of the corresponding waveform from the APG baseline and the VPG baseline. Other features based on these features were also defined, such as $b/a, c/a, d/a, e/a, (b−c−d−e)/a$, and $(b−c−d)/a$. A total of 20 VPG and APG features were defined.

Figure 2. The definition of the characteristics in PPG and its derivatives.
2.3.4. Waveform Area

The waveform AC component area features are expressed as their letters with a polyline over each. For example, the \( \hat{OS} \) feature represents the AC component area under the curve from the onset point O to the systolic peak S. A total of four waveform area features were defined.

2.3.5. Power Area

The waveform AC component power area features are expressed as their letters with a brace over each. For example, the \( \hat{\hat{OS}} \) feature represents the quadratic sum of the curve point from the onset point O to the systolic peak S. A total of 15 waveform area features were defined.

2.3.6. Ratio

The ratio features are expressed directly as their ratio formulae. For example, the \( \frac{\hat{OS}}{\hat{OO}} \) feature represents the ratio of the time span feature \( \hat{OS} \) and the time span feature \( \hat{OO} \), and the \( \frac{\hat{OS}}{\hat{OO}} \) feature represents the ratio of the curve area feature \( \hat{OS} \) and the curve area feature \( \hat{OO} \). A total of 43 ratio features were defined.

2.3.7. Slope

The slope features are expressed as their letters with a tilde over each. For example, the \( \tilde{OS} \) feature represents the slope from the onset point O to the systolic peak S. A total of 16 slope features were defined.

2.4. Feature Selection Methods

Before feature selection processing, z-score normalization was conducted for all extracted features. The mean and variance of the features were 0 and 1, respectively. The common feature selection methods included filtering, wrapping, and embedding; these methods were used to analyze the importance of the features for classification or prediction purposes from different angles. In this study, we used filtering feature selection methods [38,39], including the Spearman correlation [40], ReliefF [41], information gain (Info Gain) [39], chi-square, mRMR [42], and the Gini index. Relief is a feature selection algorithm that was first proposed for binary classification by Kira and Rendell in 1992 [41]. This algorithm chooses instances randomly and changes the weights of the relevant feature based on the nearest neighbor. Kononenko [41] proposed the ReliefF method, which is based on Relief, to solve multi-class classification. Generally speaking, ReliefF is one of the most successful strategies in feature selection. Information gain [39] is frequently used in feature selection in the field of machine learning. The information gain method regards each feature as an isolated variable and calculates its information gain; then, it evaluates the importance and relevance of each feature with a class label. The chi-square test is used to examine the dependency of the target variable (i.e., SBP) and each PPG feature; in this study, if the SBP variable was independent of each feature, then we could discard that feature variable. Conversely, if the SBP variable was dependent, then the feature variable was very important.

The minimum redundancy maximum relevance (mRMR) method [42] is a scheme that is used in feature selection to select the features most strongly correlated with a classification variable and remove the redundant features. This scheme, combined with selection features that are mutually distinct yet maintain a high correlation, comprise the selection scheme of mRMR. This combined scheme has been found to be more powerful than maximum relevance selection alone. The Gini index is a statistical measure of dispersion that is based on another statistical phenomenon, called the Lorentz curve, and it is commonly used to quantify wealth distributions, although its applications are limitless.
2.5. Classification Model

Based on the analysis of the PPG features, this study established four common linear and nonlinear classification learner models to classify the hypertension population: linear discriminant analysis (LDA), cubic SVM, weighted KNN, and logistic regression (LR).

An optimal feature subset composed of the 10 features that were selected by the feature selection method was trained and tested, and the 10-fold cross-validation was configured to evaluate the performance of the classification model. Figure 3 shows the workflow of the signal process and classification model. The data processing and modeling estimations were carried out using MATLAB software (Version R2017a, MathWorks, Inc., Natick, MA, USA).

![Flowchart of PPG signal process and hypertension classification](Figure 3. The flowchart of PPG signal process and hypertension classification.)
3. Results

3.1. Correlation Coefficient between the SBP and Features

Many studies [19,43-47] have shown that there is a linear relationship between some parameters of the PPG signal and SBP. First, the correlations between SBP and the selected features were examined. We found that several correlation coefficients were greater than 0.5 between the features and SBP. The p values of these features showed a significant difference (p < 0.01). This demonstrated that some PPG features have a significant correlation with SBP and they can therefore be used as biomarkers for hypertension. Table 1 shows a correlation coefficient list of top 10 PPG features. Figure 4 shows the regression analysis of the four PPG features with the highest correlation with SBP.

Table 1. A correlation coefficient list of top 10 features with systolic blood pressure (SBP).

| Index | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|-------|----|----|----|----|----|----|----|----|----|----|
| Feature       | (b-c)/a | b-2 d-2 | SC-2 | OS | c-2/S | S_b/c-1 | S_b/c-1 / O_a O_1 | b-2/S | SC-2 | wS_b/c-1 / O_a O_1 |
| Correlation coefficient | 0.6903 | 0.6721 | 0.6181 | 0.6164 | -0.4722 | -0.5928 | -0.6353 | -0.6391 | -0.6482 |
| p-value        | 0.0001 | 0.0009 | 0.0013 | 0.0017 | 0.0425 | 0.0058 | 0.0004 | 0.0003 | 0.0079 |

Figure 4. The regression analysis of the four features with the strongest correlation with systolic blood pressure (SBP). (A) (b-c)/a (B) b-2 d-2 (C) S_b/c-1 (D) OS.

We tested 125 features, some were examined by other researchers [33], such as OS, which is also called crest time (CT). Other, new features are defined and examined in this paper. From the features with strong correlations, we found that the most informative features lied within the interval from the b wave to the d wave. As is well-known, BP is a major physiological parameter of the human body. The features that are defined in this paper reflect the phenomenon of BP, which is important for the indirect measurement of BP and the risk stratification of the hypertensive population.
3.2. Top 10 Features Ranking

From the above correlation analysis, we found that numerous features have a significant correlation with SBP. Beyond the correlation study, we also evaluated the contribution of PPG features to hypertension classification via several feature selection methods. Table 2 shows the 10 selected features that were processed with six feature selection methods in hypertension classification. These selection methods are based on different mathematical models. Therefore, the selected features are slightly different.

Table 2. The 10 selected features processed with six feature selection methods in hypertension risk stratification, which are ranked in descending order.

| Feature Rank | Spearman | ReliefF | Info Gain | Chi2 | mRMR | Gini |
|--------------|----------|---------|-----------|------|------|------|
| 1            | b_{-2d_{-2}} | \text{S} \_1c_{-1}/O_{-1}O_{-1} | \text{S} | \text{S} \_2 | (b-c-d)/a | c_{-1}/w |
| 2            | (b-c-d)/a | \text{S} \_2 | (b-c-d)/a | c_{-1}/w | (b-c-d)/a | \text{S} \_1c_{-1}/O_{-1}O_{-1} |
| 3            | \text{S} \_2 | \text{S} \_1d_{-1}/O_{-1}O_{-1} | c_{-2}/S | \text{S} \_1c_{-1}/O_{-1}O_{-1} | c_{-1}/w | \text{S} \_2 |
| 4            | \text{S} \_2 | c_{-2}/S | (b-c-d)/a | b_{-2d_{-2}} | \text{S} \_2/\text{O} \_2 | b_{-2d_{-2}} |
| 5            | c_{-2}/S | c_{-1}/w | c_{-1}/w | (b-c-d)/a | \text{S} \_2 | \text{S} \_2 |
| 6            | \text{S} \_1c_{-1}/O_{-1}O_{-1} | (b-c-d)/a | \text{S} \_1c_{-1}/O_{-1}O_{-1} | (b-c-d)/a | b_{-2d_{-2}} | c_{-1} |
| 7            | \text{S} \_1c_{-1}/O_{-1}O_{-1} | d | b_{-2d_{-2}} | c_{-2}/S | \text{S} \_2 | \text{S} \_2 |
| 8            | b_{-2}/S | b_{-2d_{-2}} | \text{S} \_2 | c_{-2}/S | \text{S} \_1c_{-1}/O_{-1}O_{-1} |
| 9            | \text{S} \_2 | \text{S} \_2 | \text{S} \_2 | c_{-1} | \text{S} \_2 | \text{S} \_2 |
| 10           | \text{S} \_1c_{-1}/O_{-1}O_{-1} | \text{S} \_2 | c_{-1} | \text{S} \_2 | \text{S} \_2 | (b-c-d)/a |

3.3. Model Result

The optimal feature subset, which includes the top 10 features of each method selection, was produced via the various feature selection methods described above. The 10-fold cross-validation was used. For the evaluation index of risk stratification performance [48], the sensitivity (SE), positive predictive (PP), and F1-score (F1) were used to conduct the comparison and performance evaluation of the models.

Three risk stratification levels were performed: the normotensive group (48 subjects) versus the prehypertensive group (41 subjects), the normotensive group (48 subjects) versus the hypertensive group (35 subjects), and the normotensive and prehypertensive group (89 subjects) versus the hypertensive group (35 subjects). Table 3 shows the risk stratification performance results in the risk stratification process using the top 10 features.

As mentioned above, six feature selection methods were adopted to evaluate the important features. Based on the ranked features, multiple classifications were conducted using the top 10 features. As we have seen in this paper, KNN (K = 10) showed better performance in classifying the different BP categories. From Table 3, it can also be seen that KNN achieved better classification performance results than the other classifiers.

It is worth mentioning that we investigated the top 20 features as well, with the aim of improving overall accuracy. However, the classification performance results based on the top 20 features were not better than those that were based on the top 10 features. Accordingly, our conclusion was that more features did not improve the performance of the risk stratification. Therefore, in this paper, the top 10 features were analyzed and applied toward classifying the BP categories.
when the main propulsion wave reaches its peak (S). In contrast, the wave represents the smallest moment of acceleration for the main propulsion wave and it occurs. The VPG signal reflects the extent of blood volume change in the peripheral blood vessels and it occurs. The VPG signal reflects the extent of blood volume change in the peripheral blood vessels and it occurs. The PPG signal reflects physical changes of blood volume in blood vessels during cardiac. Blood pressure and a greater contribution weight with BP labels were concentrated in the interval between the systolic and diastolic periods, specifically in the interval between the b wave and the c wave. Therefore, the propagation and reflection fusion process of arterial blood during the systolic period plays an important role in the formation of SBP. The accurate analysis of this physiological process via the PPG signal is of great significance for detecting or evaluating the systolic pressure level.

The PPG signal reflects physical changes of blood volume in blood vessels during cardiac activity and blood transmission, and it is the direct representation of the blood volume state [49, 50]. The VPG signal reflects the extent of blood volume change in the peripheral blood vessels and it is the manifestation of blood change [13]. The APG signal reflects the accelerated change of blood volume, and this phenomenon largely reflects the systolic and diastolic ability of the heart. In addition, the moment of propulsion and reflection is very important for analyzing the fusion process of the propulsion wave and the reflection wave, which can be more prominent in the APG signal. Therefore, PPG, VPG, APG, and the third derivative can reflect cardiac activity in the vascular circulatory system from different dimensions, and therefore the systematic and coherent analysis of waveform signals is very important in understanding the physiological process more clearly.

4. Discussion

We know that human blood vessels and microcirculation systems undergo changes with increases to BP. These changes are especially obvious for patients with severe hypertension. PPG signals contain abundant information that can reflect the change process. Therefore, high-quality signal acquisition of PPG signals is very important for accurately assessing the intrinsic relationship between BP and PPG signals.

The SBP of human arteries is the maximum value that is formed in the middle of the systolic period under the action of the fusion of blood propulsion and reflection from the heart. Therefore, the propulsion and reflection fusion process of arterial blood during the systolic period plays an important role in the formation of SBP. The accurate analysis of this physiological process via the PPG signal is of great significance for detecting or evaluating the systolic pressure level.

The PPG signal reflects physical changes of blood volume in blood vessels during cardiac activity and blood transmission, and it is the direct representation of the blood volume state [49, 50]. The VPG signal reflects the extent of blood volume change in the peripheral blood vessels and it is the manifestation of blood change [13]. The APG signal reflects the accelerated change of blood volume, and this phenomenon largely reflects the systolic and diastolic ability of the heart. In addition, the moment of propulsion and reflection is very important for analyzing the fusion process of the propulsion wave and the reflection wave, which can be more prominent in the APG signal. Therefore, PPG, VPG, APG, and the third derivative can reflect cardiac activity in the vascular circulatory system from different dimensions, and therefore the systematic and coherent analysis of waveform signals is very important in understanding the physiological process more clearly.

From the feature selection analysis, we also found that features with a strong correlation to systolic pressure and a greater contribution weight with BP labels were concentrated in the interval between the systolic and diastolic periods, specifically in the interval between the b wave and the c wave. The b wave represents the smallest moment of acceleration for the main propulsion wave and it occurs when the main propulsion wave reaches its peak (S). In contrast, the c wave is the largest moment of acceleration for the reflected wave. This interval shows significant differences in the PPG waveform due to different BP levels. Therefore, the study and analysis of the physiological changes of blood vessels in the b wave and c wave periods can help to clarify the formation of the systolic pressure peak more accurately. Most importantly, the precise comprehension of this process is very useful for
improving the accuracy of systolic pressure detection. In addition, signal sampling accuracy needs special attention. Generally, the interval between the \( b \) wave and \( c \) wave is about 60–130 ms (median: 90 ms) and it accounts for 7–14% (median: 11%) of a heartbeat cycle. Therefore, low signal sampling accuracy loses a substantial amount of important information. To avoid any under-sampling issues, such as ignoring critical PPG features in the data collection phase, we used a high sampling frequency rate of 1 kHz with a high AD conversion rate of 12 bits.

In this study, some features reflected the details of the fusion process of the main propulsion wave and the reflected wave, such as \( \frac{S c_{-2}}{S} \), \( \frac{c_{-2}}{c_{-1}} / O_{+1} O_{+1} \), and \( b_{-2} d_{-2} \). In Table 1, the top 10 features include two time span features, two slope features, three power areas, and three amplitude ratio features. Seven features’ correlation coefficients of the top 10 features were greater than 0.6. More importantly, the features are mainly related with the \( b \), \( S \), \( c \), and \( d \) waves. As is well-known, the \( S \) wave and \( c \) wave exhibit different morphological changes under different physiological statuses, especially different BP levels [13,30]. Therefore, the selected features in this study are good expressions of the conditions of the main propulsion wave and reflected wave. This result confirms our findings on the MIMIC database, which indicate that the \( b d \) area is associated with hypertension [28].

We know that the moment of the fusion of the reflection wave and the main propulsion wave has a decisive effect on the actual SBP. The early fusion of the reflection wave and the main propulsion wave leads to an increase in human BP, while delayed fusion generates adverse effects on the maintenance of the ideal level (90–120 mmHg/60–80 mmHg) of BP. The features in this study have clearer physiological explanations than \( a \), \( b \), \( c \), \( d \), \( e \), and \( \frac{(b–c–d–e)}{a} \). The features of \( a \), \( b \), \( c \), \( d \), and \( e \), and their combination, reflect the ability of blood propulsion in the blood vessels, which can in turn reflect the physiological process of BP formation.

However, the features in this study directly reflect the formation of BP under the influence of multiple factors, and they are thus closer to the physiological significance of BP; therefore, they serve as the theoretical and physiological basis for the realization of non-cuff continuous ambulatory BP measurement. The characteristic expressions of physiological phenomena in the cardiovascular circulation process are of great significance to the development of non-invasive BP detection techniques and the non-invasive evaluation of cardiovascular aging and other physiological phenomena. Note that the approach achieves the six-step biomedical signal analysis framework for tackling non-communicable diseases, such as hypertension [51].

In Table 2, most of the useful features were selected by different feature selection methods, which are based on different theories. Here, ReliefF combined with the KNN classifier achieved better risk stratification performance results than the other models. ReliefF is an excellent feature weighting method for the classification problem. Less relevant features are removed, while more relevant features are retained, and all of the features are evaluated one by one, with the selected features representing those that are most useful for classification. In addition, as was shown earlier, the experimental results of the normotensive group versus the hypertensive group were better than those of the normotensive group versus the prehypertensive group and the normotensive and prehypertensive groups versus the hypertensive group.

We know that in the treatment and management of hypertension [4,52], doctors generally do not provide drug or medical treatment but instead advise patients to adjust their diets and lifestyles to restore normal BP levels. In some patients, diet and lifestyle changes are enough, and medical therapy may not be necessary, especially in prehypertensive or stage 1 hypertensive patients [7]. However, medication is needed to sufficiently reduce BP for most stage 1 and almost all stage 2 hypertension cases [7]. Stage 2 hypertension cases also require longer-term treatment and drug control than stage 1 hypertension cases [7].

Hypertension in different stages can lead to varying degrees of cardiovascular circulation change; some changes are mild or reversible (unstable changes, such as prehypertension or stage 1 hypertension), while others are qualitative (chronic changes, such as stage 2 hypertension and severe
hypertension, for which it is difficult to restore the normal BP level). Therefore, the cardiovascular circulatory system will exhibit different reflections for different BP levels [53].

One limitation of our study was that we examined the correlation of PPG features with only SBP. Finding the optimal features for detecting DBP and mean arterial pressure is our next step. Moreover, we must still investigate the classification between prehypertension and hypertension.

For chronic changes in the cardiovascular system, the risk stratification can show higher performance; in contrast, for reversible, unstable changes, which are prone to fluctuations in BP, the risk stratification shows relatively low performance. This result also illustrates the significance of long-term, dynamic BP monitoring. Yet, there are some limits to this approach. In this study, the sampling rate of the device was configured as 1 kHz. A high sampling rate and greater sampling precision consume more power, and therefore the sampling frequency must be reduced to save power so that the device can collect data for a longer period of time. As is known, the power consumption of wearable or portable devices is high. Therefore, low power consumption and a more efficient algorithm design must be achieved. In addition, the dataset was collected from an experimental room in which noise and other types of interference were avoided as much as possible. Thus, future work should be conducted in real-world conditions.

5. Conclusions

In this paper, we discussed hypertension risk stratification that was based on the morphological features of PPG signals and machine learning technology. By analyzing the physiological significance of the PPG features, the quantitative expression of SBP formation was realized, and the intrinsic relation between SBP formation and PPG features was established. By using six feature selection methods, 10 PPG features were found to be associated with systolic pressure. We observed the effects of different BP levels on physical changes in the cardiovascular circulatory system in the hypertension classification experiments with the optimal feature subsets, and these changes were more directly reflected in the PPG signal and its derivative signals. The risk stratification model showed higher performance in detecting severe hypertension than non-severe hypertension (such as prehypertension). In the PPG signal, a certain variability and instability were also observed. Therefore, higher accuracy in long-term, dynamic monitoring and evaluation for detection or screening is more realistic. The method that is described in this paper can hopefully achieve ambulatory BP measurement based only on the PPG signal. In future studies, we will develop and examine the proposed method for long-term monitoring of BP.

Author Contributions: Y.L. and M.E. performed the statistical analysis. Y.L., Z.C., R.W., and M.E. conceived the study and drafted the manuscript. M.E. and Z.C. designed the study and led this investigation. All authors approved the final manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (grant No. 61627807, 81873913), the study abroad program for graduate student of Guilin University of Electronic Technology and Promotion of Basic Ability of Young and Middle-aged Teachers in Universities of Guangxi (2018KY0217). The APC was funded by the Natural Science Foundation of Guangxi (2017GXNSFGA198005).

Acknowledgments: We thank Shaozhen Tan and Tanqi Zhou’s help in the experimental preparation stage. And we also thank Shimin Yin and Daming Su’s in the development stage of device and its app. They provided many suggestions and their knowledge to make the device, app and experiment perfect.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gepner, A.D.; Tedla, Y.; Colangelo, L.A.; Tattersall, M.C.; Korcarz, C.E.; Kaufman, J.D.; Liu, K.; Burke, G.L.; Shea, S.; Greenland, P.; et al. Progression of Carotid Arterial Stiffness with Treatment of Hypertension Over 10 Years: The Multi-Ethnic Study of Atherosclerosis. *Hypertension* 2017, 69, 87–95. [CrossRef] [PubMed]
2. Alhamdow, A.; Lindh, C.; Albin, M.; Gustavsson, P.; Tinnerberg, H.; Broberg, K. Early markers of cardiovascular disease are associated with occupational exposure to polycyclic aromatic hydrocarbons. *Sci. Rep.* 2017, 7, 9426. [CrossRef] [PubMed]

3. Leung, A.A.; Daskalopoulou, S.S.; Dasgupta, K.; McBrien, K.; Butalia, S.; Zarrke, K.B.; Nerenberg, K.; Harris, K.C.; Nakhla, M.; Cloutier, L.; et al. Hypertension Canada’s 2017 Guidelines for Diagnosis, Risk Assessment, Prevention, and Treatment of Hypertension in Adults. *Can. J. Cardiol.* 2017, 33, 557–576. [CrossRef] [PubMed]

4. Drawz, P.E.; Pajewski, N.M.; Bates, J.T.; Bello, N.A.; Cushman, W.C.; Dwyer, J.P.; Fine, L.J.; Goff, D.C., Jr.; Haley, W.E.; Krousel-Wood, M.; et al. Effect of Intensive Versus Standard Clinic-Based Hypertension Management on Ambulatory Blood Pressure: Results from the SPRINT (Systolic Blood Pressure Intervention Trial) Ambulatory Blood Pressure Study. *Hypertension* 2017, 69, 42–50. [CrossRef]

5. Kachuee, M.; Kiani, M.M.; Mohammadzade, H.; Shabany, M. Cuffless Blood Pressure Estimation Algorithms for Continuous Health-Care Monitoring. *IEEE Trans. Biomed. Eng.* 2017, 64, 859–869. [CrossRef]

6. Atomi, K.; Kawanaka, H.; Bhuiyan, M.S.; Oguri, K. Cuffless Blood Pressure Estimation Based on Data-Oriented Continuous Health Monitoring System. *Comput. Math. Methods Med.* 2017, 2017, 1803485. [CrossRef] [PubMed]

7. Chobanian, A.V.; Bakris, G.L.; Black, H.R.; Cushman, W.C.; Green, L.A.; Izzo, J.L., Jr.; Jones, D.W.; Materson, B.J.; Oparil, S.; Wright, J.T., Jr.; et al. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure. *Hypertension* 2003, 42, 1206–1252. [CrossRef]

8. Watanabe, N.; Bando, Y.K.; Kawachi, T.; Yamakita, H.; Futatsuyama, K.; Honda, Y.; Yasui, H.; Nishimura, K.; Kamihara, T.; Okumura, T.; et al. Development and Validation of a Novel Cuff-Less Blood Pressure Monitoring Device. *JACC Basic Transl. Sci.* 2017, 2, 631–642. [CrossRef]

9. Korotkoff, N.S. On methods of studying blood pressure. *Bull. Imp. Mil. Med. Acad.* 1905, 11, 365–367.

10. Yan, Y.S.; Zhang, Y.T. Noninvasive estimation of blood pressure using photoplethysmographic signals in the period domain. In *Proceedings of the Annual International Conference*, Shanghai, China, 17–18 January 2005; Volume 4, pp. 3583–3584. [CrossRef]

11. Hui, X.; Kan, E.C. Monitoring vital signs over multiplexed radio by near-field coherent sensing. *Nat. Electron.* 2017. [CrossRef]

12. Hosseini, Z.S.; Zahedi, E.; Movahedian Attar, H.; Fakhrzadeh, H.; Parsa, M.H. Discrimination between different degrees of coronary artery disease using time-domain features of the finger photoplethysmogram in response to reactive hyperemia. *Biomed. Signal Process. Control* 2015, 18, 282–292. [CrossRef]

13. Elgendi, M. On the analysis of fingertip photoplethysmogram signals. *Curr. Cardiol. Rev.* 2012, 8, 14–25. [CrossRef]

14. Elgendi, M.; Fletcher, R.R.; Norton, I.; Brearley, M.; Abbott, D.; Lovell, N.H.; Schuurmans, D. Frequency analysis of photoplethysmogram and its derivatives. *Comput. Methods Prog. Biomed.* 2015, 122, 503–512. [CrossRef] [PubMed]

15. Hughes, T.M.; Craft, S.; Lopez, O.L. Review of ‘the potential role of arterial stiffness in the pathogenesis of Alzheimer’s disease’. *Neurodegener. Dis. Manag.* 2015, 5, 121–135. [CrossRef] [PubMed]

16. Mukkamala, R.; Hahn, J.O.; Inan, O.T.; Mestha, L.K.; Kim, C.S.; Toreyin, H.; Kyal, S. Toward Ubiquitous Blood Pressure Monitoring via Pulse Transit Time: Theory and Practice. *IEEE Trans. Bio-Med. Eng.* 2015, 62, 1879–1901. [CrossRef]

17. Palmeri, L.; Gradwohl, G.; Nitzan, M.; Hoffman, E.; Adar, Y.; Shapir, Y.; Koppel, R. Photoplethysmographic waveform characteristics of newborns with coarctation of the aorta. *J. Perinatol.* 2017, 37, 77–80. [CrossRef]

18. Allen, J.; Murray, A. Age-related changes in peripheral pulse timing characteristics at the ears, fingers and toes. *J. Hum. Hypertens.* 2002, 16, 711–717. [CrossRef]

19. Shin, H.; Min, S.D. Feasibility study for the non-invasive blood pressure estimation based on ppg morphology: Normotensive subject study. *Biomed. Eng. Online* 2017, 16, 10. [CrossRef]

20. Alian, A.A.; Shelley, K.H. Photoplethysmography. Best practice & research. *Clin. Anaesthesiol.* 2014, 28, 395–406. [CrossRef]

21. Nichols, W.; O’Rourke, M.; Vlachopoulos, C. (Eds.) *McDonald’s Blood Flow in Arteries: Theoretical*; CRC Press: Boca Raton, FL, USA, 2011.
22. Tang, S.C.; Huang, P.W.; Hung, C.S.; Shan, S.M.; Lin, Y.H.; Shieh, J.S.; Lai, D.M.; Wu, A.Y.; Jeng, J.S. Identification of Atrial Fibrillation by Quantitative Analyses of Fingertip Photoplethysmogram. *Sci. Rep.* 2017, 7, 45644. [CrossRef]

23. Koivistoinen, T.; Lyytikäinen, L.P.; Aatola, H.; Luukkaala, T.; Juonala, M.; Viikari, J.; Lehtimäki, T.; Raitakari, O.T.; Kähönen, M.; Hutri-Kähönen, N. Pulse Wave Velocity Predicts the Progression of Blood Pressure and Development of Hypertension in Young Adults. *Hypertension* 2018. [CrossRef] [PubMed]

24. Ding, X.; Yan, B.P.; Zhang, Y.T.; Liu, J.; Zhao, N.; Tsang, H.K. Pulse Transit Time Based Continuous Cuffless Blood Pressure Estimation: A New Extension and A Comprehensive Evaluation. *Sci. Rep.* 2017, 7, 11554. [CrossRef] [PubMed]

25. Elgendi, M.; Fletcher, R.; Norton, I.; Brearley, M.; Abbott, D.; Lovell, N.H.; Schuurmans, D. On Time Domain Analysis of Photoplethysmogram Signals for Monitoring Heat Stress. *Sensors* 2015, 15, 24716. [CrossRef] [PubMed]

26. Liang, Y.; Chen, Z.; Liu, G.; Elgendi, M. A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China. *Sci. Data* 2018, 5, 180020. [CrossRef] [PubMed]

27. Liang, Y.; Elgendi, M.; Chen, Z.; Ward, R. An optimal filter for short photoplethysmogram signals. *Sci. Data* 2018, 5, 180076. [CrossRef] [PubMed]

28. Liang, Y.; Chen, Z.; Ward, R.; Elgendi, M. Hypertension Assessment via ECG and PPG Signals: An Evaluation Using MIMIC Database. *Diagnostics* 2018, 8, 65. [CrossRef] [PubMed]

29. Martínez, G.; Howard, N.; Abbott, D.; Lim, K.; Ward, R.; Elgendi, M. Can Photoplethysmography Replace Arterial Blood Pressure in the Assessment of Blood Pressure? *J. Clin. Med.* 2018, 7, 316. [CrossRef] [PubMed]

30. Elgendi, M.; Liang, Y.; Ward, R. Toward Generating More Diagnostic Features from Photoplethysmogram Waveforms. *Diseases* 2018, 6, 20. [CrossRef]

31. Liang, Y.; Chen, Z.; Ward, R.; Elgendi, M. Photoplethysmography and Deep Learning: Enhancing Hypertension Risk Stratification. *Biosensors* 2018, 8, 101. [CrossRef]

32. Elgendi, M. Optimal Signal Quality Index for Photoplethysmogram Signals. *Bioengineering* 2016, 3, 21. [CrossRef]

33. Suganthi, L.; Manivannan, M.; Kunwar, B.K.; Joseph, G.; Danda, D. Morphological analysis of peripheral arterial signals in Takayasu’s arteritis. *J. Clin. Monit. Comput.* 2015, 29, 87–95. [CrossRef] [PubMed]

34. Elgendi, M. TERMA Framework for Biomedical Signal Analysis: An Economic-Inspired Approach. *Biosensors* 2016, 6, 55. [CrossRef] [PubMed]

35. Elgendi, M. Eventogram: A Visual Representation of Main Events in Biomedical Signals. *Bioengineering* 2016, 3, 22. [CrossRef]

36. Elgendi, M. Detection of c, d, and e waves in the acceleration photoplethysmogram. *Comput. Methods Prog. Biomed.* 2014, 117, 125–136. [CrossRef]

37. Elgendi, M.; Norton, I.; Brearley, M.; Abbott, D.; Schuurmans, D. Detection of a and b waves in the acceleration photoplethysmogram. *Biomed. Eng. Online* 2014, 13, 139. [CrossRef] [PubMed]

38. Zhao, Z.; Morstatter, F.; Sharma, S.; Alelyani, S.; Anand, A.; Liu, H. Advancing Feature Selection Research; ASU: Tempe, AZ, USA, 2010; pp. 1–28.

39. Alhaj, T.A.; Siraj, M.M.; Zainal, A.; Elshoush, H.T.; Elhaj, F. Feature Selection Using Information Gain for Improved Structural-Based Alert Correlation. *PLoS ONE* 2016, 11, e0166017. [CrossRef]

40. Hauke, J.; Kossowski, T. Comparison of Values of Pearson’s and Spearman’s Correlation Coefficients on the Same Sets of Data. *Quaest. Geogr.* 2011, 30. [CrossRef]

41. Robnik-Šikonja, M.; Kononenko, I. Theoretical and empirical analysis of ReliefF and RReliefF. *Mach. Learn.* 2003, 53, 23–69. [CrossRef]

42. Yu, H. Efficient Feature Selection via Analysis of Relevance and Redundancy. *J. Mach. Learn. Res.* 2004, 5, 1205–1224.

43. Lin, W.-H.; Wang, H.; Samuel, O.W.; Liu, G.; Huang, Z.; Li, G. New photoplethysmogram indicators for improving cuffless and continuous blood pressure estimation accuracy. *Physiol. Meas.* 2018, 39, 025005. [CrossRef]

44. Hsiu, H.; Hsu, C.L.; Wu, T.L. Effects of different contacting pressure on the transfer function between finger photoplethysmographic and radial blood pressure waveforms. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* 2011, 225, 575–583. [CrossRef] [PubMed]
45. Jeong, I.; Jun, S.; Um, D.; Oh, J.; Yoon, H. Non-invasive estimation of systolic blood pressure and diastolic blood pressure using photoplethysmograph components. *Yonsei Med. J.* 2010, 51, 345–353. [CrossRef]

46. Li, Y.; Wang, Z.; Zhang, L.; Yang, X.; Song, J. Characters available in photoplethysmogram for blood pressure estimation: Beyond the pulse transit time. *Aust. Phys. Eng. Sci. Med.* 2014, 37, 367–376. [CrossRef] [PubMed]

47. Kim, J.S.; Kim, K.K.; Baek, H.J.; Park, K.S. Effect of confounding factors on blood pressure estimation using pulse arrival time. *Physiol. Meas.* 2008, 29, 615–624. [CrossRef] [PubMed]

48. Pombo, N.; Garcia, N.; Bousson, K. Classification techniques on computerized systems to predict and/or to detect Apnea: A systematic review. *Comput. Methods Prog. Biomed.* 2017, 140, 265–274. [CrossRef]

49. Allen, J. Photoplethysmography and its application in clinical physiological measurement. *Physiol. Meas.* 2007, 28, R1–R39. [CrossRef]

50. Allen, J.; Murray, A. Age-related changes in the characteristics of the photoplethysmographic pulse shape at various body sites. *Physiol. Meas.* 2003, 24, 297–307. [CrossRef]

51. Elgendi, M.; Howard, N.; Lovell, N.; Cichocki, A.; Bearley, M.; Abbott, D.; Adatia, I. A Six-Step Framework on Biomedical Signal Analysis for Tackling Noncommunicable Diseases: Current and Future Perspectives. *JMIR Biomed. Eng.* 2016, 1, e1. [CrossRef]

52. Gudsoorkar, P.S.; Tobe, S.W. Changing concepts in hypertension management. *J. Hum. Hypertens.* 2017. [CrossRef]

53. Bruno, R.M.; Duranti, E.; Ippolito, C.; Segnani, C.; Bernardini, N.; Di Candio, G.; Chiarugi, M.; Taddei, S.; Virdis, A. Different Impact of Essential Hypertension on Structural and Functional Age-Related Vascular Changes. *Hypertension* 2017, 69, 71–78. [CrossRef]

© 2018 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 (http://creativecommons.org/licenses/by/4.0/).