ABSTRACT

Regarding heart failure (HF), reducing mortality and prolonging life is one of the main treatment goals. Many clinical studies define HF patients according to Left Ventricular Ejection Fraction (LVEF). Two different subtypes in patients with HF are: HF with preserved ejection fraction (HFrEF) and HF with reduced ejection fraction (HFpEF). Echocardiography is generally used to measure LVEF. This is a pricy device that requires an expert and there may be situations where attaining the device is restricted. There may be cases that treatment should be started without echocardiography. Economical and practical measurement and decision support systems are needed to solve such situations. In this study, an algorithm was improved to detect HFrEF and HFpEF by using solely heart rate variability (HRV) derived from photoplethysmography (PPG). PPG data were obtained from volunteers for 10s, digital filters were used to clean PPGs, and HRV derivation was made from cleaned PPG. Totally thirty-seven features were obtained. Consequently, features were selected, and classification that was realized with only 3 features extracted from HRV gave significant results. 10-fold cross validation was performed for evaluation. The classification performance parameters were: accuracy %98.33, sensitivity 0.967, specificity 1, AUC 0.983, F-measure 0.981 and Kappa 0.967. This study provided highly reliable non-random results for distinguishing between HFrEF and HFpEF. This system, which works with such high performance with traditional machine learning methods used in real-time systems, makes a significant contribution to the literature in terms of diagnosing HFrEF and HFpEF cases with a single signal.

Keywords: Machine learning, Artificial intelligence, PPG-derived HRV, HFrEF, HFpEF, Classification.

1. INTRODUCTION

Ejection fraction (EF) represents the pumping efficiency of the heart. It is related to the blood amount that is pumped from a ventricle with a heartbeat. It is the ratio of the pulse volume to the end-diastolic volume. The efficiency of ejecting to the circulation of systemic is Left Ventricular EF (LVEF), and the efficiency of ejecting to the circulation of pulmonary is the right ventricular EF. Echocardiography
is the gold standard to measure EF. LVEF is generally low in systolic HF. It is a significant determinant of the prognosis of systolic HF. The heart rate may change (low or high) in a healthy person during the daily course, but reduced LVEF is evermore an indicator of HF [1].

Structural or functional disorganization resulting in the heart cannot supply enough oxygen to the tissues can be used for HF definition [2]. HF is described as a symptom complex that results from a structural or functional failure in the heart, and HF patients have signs like dyspnoea, ankle swelling, and fatigue and symptoms like apex beat shift, pulmonary crepitation, and elevated jugular vein pressure [3].

When systolic impairment rises, LVEF progressively reduces, end-systolic volumes generally rise. Many clinical studies define HF patients according to LVEF, so it is more significant than only an indication of HF.

At present attested treatments are efficient just in HFrEF or systolic HF. Furthermore, studies were also performed about HF patients of which LVEF is higher than 40-45% and have no other relating to the heart defects. LVEF was quite typical (generally higher than 50%) in some of them and no important decrease was seen in systolic function. Thus, to identify these patients, "HF with preserved EF (HFpEF)" term was improved [4].

In HFpEF patients, the underlying pathophysiological defect is considered to become LV diastolic dysfunction. So, for diagnosing HFpEF, LV diastolic dysfunction needs to be diagnosed. Any sole parameter of echocardiography is not adequately definitive and repeatable for defining LV diastolic disturbance. Thus, an exhaustive echocardiographic consultation with completely correlated bidimensional Doppler data is advised. Devices for Doppler ultrasound are pricy, and an expert is necessary for the echocardiographic consultation which is laborious. Since HFpEF might originate from structural and functional disturbances relating to the heart instead of only one malady presence, it is a sophisticated syndrome, even HF professionals have trouble making accurate diagnoses [5].

Physiological indicators like the amount of oxygen in the blood and cardiac outflow are widely measured using photoplethysmography (PPG). PPG is a well choice for health studies relating to the blood vessels and heart owing to the little dimension and cost-effectiveness. Moreover, researches demonstrate that this signal could be acquired regardless of calibration [6].

There were many studies using PPG, in most of which machine learning (ML) algorithms were used, on cardiovascular (CV) and heart diseases. Groups of CV diseased and healthy individuals were appreciated with a ML algorithm using synchronous ear and finger PPG signals [7]. Blood pressure prediction study was non-invasively made using PPG with ML and artificial intelligence (AI) methods [8].

For alternating to invasive techniques, studies have also been carried out using PPG. It was ascertained to indicate LV filling pressure in the course of the Valsalva maneuver [9]. The benefit of PPG in the course of the Valsalva Maneuver to specify whether HF patients admitted to the hospital for treatment were at risk was studied [10].
As PPG is measured easily, there are many studies on electronic devices, wearables, and smartphones. Using PPG, atrial fibrillation detection was made with a new algorithm using smartphones [11]. As PPG is easy to measure from wearables and smartphones, a study was made to detect diabetes using PPG [12]. A wristlet with PPG technology that could determine the heart rate and RR apertures with significant accuracy was studied [13].

An electronic device using PPG was improved for distinguishing congestive heart failure (CHF) patients and healthy persons [14]. A distinction between healthy and CHF patients was obtained using simultaneously collected ECG and PPG [15].

Studies on the HFrEF and HFpEF subtypes were also conducted. Neurophysiological differences of HF subtypes (HFrEF and HFpEF) in Nigeria were investigated. In the study, cognitive tests with universal validity were applied to both subtypes of patients. Echocardiographic and clinical correlation analysis with cognitive performance was made [16]. Using features such as comorbidity accumulation rates, frequency of hospitalization, and differences in the use of special care for both HF subtypes, it was determined that these subtypes showed different clinical comorbidities and disease progression patterns [17]. A study was performed to estimate HFrEF and HFpEF using features such as age, gender, body mass index, systolic blood pressure, antihypertensive therapy, and past myocardial infarction [18]. Using older age, history of diabetes and valvular disease, body mass index, smoking, and atrial fibrillation as features, risk prediction was made for HFpEF and HFrEF subtypes [19]. Different machine learning methods that predict HFrEF and HFpEF were compared using demographics such as gender, age, and features obtained by invasive methods such as hemoglobin, sodium, and glucose [20].

CHF contains the HFpEF and HFrEF cases. In the literature, studies on CHF are usually concerned about HF presence or absence. Features obtained by demographic, echocardiographic, or invasive methods were used in the binary classification studies of HFrEF and HFpEF. There is no study in the literature in which only the PPG signal is used. This study stands out in terms of distinguishing the cases of HFrEF and HFpEF with only the easily obtainable PPG signal. In this study, HFpEF and HFrEF distinction was made using PPG with machine learning methods.

Because HF patients have some similar symptoms as in chest diseases, such as fatigue, chest ache, and dyspnoea, particularly HFpEF cases (LVEF ≥ 50% as in healthy persons) may be steered to a pulmonologist. If the chest disease expert is cautious, B-type natriuretic peptide (BNP) and N-terminal proBNP (NT-proBNP) tests, which are blood tests, are done and the cases are defined as HFpEF. However, when solely chest disease tests are done and there is no abnormality, the HFpEF case may be overlooked. BNP and NT-proBNP are invasive and troublesome techniques. Furthermore, echocardiography is a pricy device, requires expert, and cases might be where attaining the device is restricted. Thus, for alternating to such states of affairs, economical and practical diagnosis technique was researched.

With this study, it is aimed to design a system that can diagnose HFpEF and HFrEF cases, which are difficult to diagnose even for experts and can be overlooked, with a single signal.
In the literature, Heart Rate Variability (HRV) was used in studies about heart failure [21]. Albeit HRV was usually derived from ECG, HRVs derived from PPG were also used [22]. So, we thought that PPG-derived HRVs could be used for our study.

Statistical methods were used to analyze the difference between classes for investigating the usefulness of the HRV's features that could be utilized for the diagnosis algorithm in the absence of echocardiography, demographics, and other exhaustive tests.

The most distinguishing features were detected. The features were abated by selecting them, then studies about classification were achieved with ML algorithms. The best classifier results were acquired by Support Vector Machines (SVMs) technique.

2. MATERIALS and METHODS

Some of the data was acquired from the patients who came to the cardiology outpatient clinic of the Training and Research Hospital of Sakarya University and some of them were from the patients who were hospitalized in the cardiology service. The ethics committee report numbered 16214662/050.01.04/123 was acquired from the Dean's Office of Medicine Faculty of Sakarya University. Data collecting and data recording were performed with a Biopac MP36 device. Echocardiographic outcomes of volunteers that could be included in this research were evaluated by an expert cardiologist and the data were acquired. PPGs were obtained from the right signing fingers of sixty-one individuals (twenty-five years old or older) for 10 s. 200 Hz was the sampling frequency. PPG data from each volunteer were received during 10 s. PPGs from the same individuals at several time apertures were also employed. The dataset included 120 data, 60 of which were HFrEF, and 60 of which were HFpEF. Volunteers' demographic information was given in Table 1.

Table 1. Volunteers' Demographics.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Height (cm)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
<th>Age (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean ± SD for females</td>
<td>67.2 ± 12.32</td>
<td>75.75 ± 10.93</td>
<td>172.6 ± 5.98</td>
<td>23.38 ± 2.94</td>
</tr>
<tr>
<td>Mean ± SD for males</td>
<td>72.18 ± 12.81</td>
<td>69.09 ± 29.27</td>
<td>158.18 ± 6.03</td>
<td>27.56 ± 11.3</td>
</tr>
<tr>
<td>Mean ± SD for total</td>
<td>68.97 ± 12.52</td>
<td>73.39 ± 19.28</td>
<td>167.48 ± 9.16</td>
<td>26.15 ± 7.01</td>
</tr>
</tbody>
</table>

2.1. Data Preprocessing

Figure 1 indicates the study's workflow diagram. This study was performed based on this flow and the outcomes were acquired. Following data acquisition, the data labelling was made by a specialist cardiologist. Firstly, the PPG signals were cleaned from artifacts and noise. The signal was filtered by Chebyshev Type II band-pass filter in the band of 0.25 – 100 Hz, stopband attenuation was stated as 60 dB. For the mains noise which is at 50 Hz, the signal was filtered by a notch filter in the interval of 49-51 Hz. Stopband attenuation was also 60 dB. As final, the signal was filtered with Moving Average Filter. Afterwards, from cleaned PPGs, HRVs were derived.
Sample PPGs of patients with HFrEF and HFpEF was presented in Figure 2. Further, Fast Fourier Transform was used to obtain the periodogram graph of signals and this graph was also presented in the figure. Whilst there was a similarity between the PPGs, they were distinguished in the periodogram. Hence, periodogram usage is advantageous for ML methods. After, features were extracted from HRVs. The features were statistically analyzed and they were selected. Eventually, we made classification studies with the features selected and we developed the diagnostic algorithm.
2.2. HRV Features
Firstly, 21 features in the time domain were extracted from the HRVs. Table 2 shows these features [23]. In the 1st column the numbers given to the features, in the 2nd column the features' names and in the 3rd column the formulas of the features were given in Table 2. Furthermore, sixteen more features were also attached to the twenty-one features. Of the sixteen features were eight of which the output parameters of Burg's Method and eight of which the output parameters of the Yule-Walker Method. Each of these eight output parameters were four of which were the normalized autoregressive (AR) parameters equivalent to the fourth-order model, one of which was the predicted variance of the white noise input, and three of which were the reflection coefficients. The whole number of features extracted from HRV was 37. MATLAB was used to calculate features.

2.3. Statistical Analysis
The distribution of the HRV is not normal. For this reason, statistical analysis is made with non-parametric tests.

The Mann Whitney U test is performed to establish whether the two groups belong to the same population [24]. In this study, the Mann Whitney U test was performed to define which of the 37 extracted features were distinctive. Asymptotic Significances (p-values) of the features were given in Table 3. The features were also ranked in the table.

Table 2. Features Extracted From HRV at Time Domain.

<table>
<thead>
<tr>
<th>Feature Number</th>
<th>Feature Name</th>
<th>Feature Equation</th>
</tr>
</thead>
</table>
1. **Kurtosis**

\[ x_{kur} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{(n-1)S^4} \]

2. **Skewness**

\[ x_{ske} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(n-1)S^3} \]

3. **Interquartile range**

\[ IQR = Iqr(x) \]

4. **Geometric mean**

\[ \text{GeoM} = \sqrt[n]{x_1 \cdots x_n} \]

5. **Harmonic mean**

\[ \text{HarM} = \frac{n}{\left(\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right)} \]

6. **Hjort Parameters - Mobility**

\[ M = \frac{S_1^2}{S_2^2} \]

7. **Hjort Parameters - Complexity**

\[ C = \sqrt{(S_2^2/S_1^2)^2 - (S_1^2/S_2^2)^2} \]

8. **Hjort Parameters - Activity**

\[ A = S^2 \]

9. **Maximum**

\[ x_{max} = \max(x_i) \]

10. **Median**

\[ \bar{x} = \begin{cases} \frac{x_{n+1}}{2} & : x \text{ odd} \\ \frac{1}{2} \left( \frac{x_n + x_{n+1}}{2} \right) & : x \text{ even} \end{cases} \]

11. **Mean Absolute Deviation**

\[ \text{MAD} = mad(x) \]

12. **Minimum**

\[ x_{min} = \min(x_i) \]

13. **Central Moments**

\[ \text{CM} = \text{moment}(x, 10) \]

14. **Mean**

\[ \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i \]

15. **Average curve length**

\[ CL = \frac{1}{n} \sum_{i=2}^{n} |x_i - x_{i-1}| \]

16. **Standard Deviation**

\[ S = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}) \]

17. **Standard Error**

\[ S_x = S/\sqrt{n} \]

18. **Shape Factor**

\[ \text{SF} = X_{rms} / \left( \frac{1}{n} \sum_{i=1}^{n} \sqrt{|x_i|} \right) \]

19. **Singular Value Decomposition**

\[ \text{SVD} = \text{svd}(x) \]

20. **25% Trimmed Mean**

\[ \text{TM25} = \text{trimmean}(x, 25) \]

21. **50% Trimmed Mean**

\[ \text{TM50} = \text{trimmean}(x, 50) \]

* MATLAB was used to calculate the feature
\[ \bar{x}: \text{the distribution mean} \]
\[ n: \text{the sample’s observations number} \]
\[ S_1: \text{the standard deviation of the signal’s 1st derivative} \]
\[ S_2: \text{the standard deviation of the signal’s 2nd derivative} \]
\[ X_{rms}: \text{Root Mean Square} \]
\[ X_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} |x_i|^2} \]
According to Table 3, \( p \)-value is higher than 0.05 for the features 3, 6, 9, 17, 22, 23, 24, 25, 26, 27, 30, 31, 32, 33, 36, 37. That is, these 16 features are not distinctive for the two classes.

Classification with high accuracy was made with only three features as 10th, 20th, and 21st. These features were in the top three in the ranking.

2.4. Study of Classification

The dataset was arranged as a matrix. Rows were individuals and columns were features. After labeling, a label column was attached next to the feature columns as the last column. 10-fold cross validation was performed for evaluation. Performance values were calculated by comparing the pre-arranged label column and simulation results.

Inclusive of Support Vector Machines (SVMs), Decision Trees, and k-Nearest Neighbours (k-NN), three ML algorithms were applied.

Fine Gaussian SVM was applied. Kernel function was Gaussian, kernel scale was 0.35, and box constraint was 1. Fine Decision Trees was applied. Maximum number of splits was 100, and split criterion was Gini's diversity index. And for the fine k-NN, number of neighbors was 1, distance metric was Euclidean, and distance weight was equal.
As a result of classifications made with only 3 features of HRV as median, 25% trimmed mean and 50% trimmed mean, very high and reliable results were obtained.

For each classifier, performance parameters of classification were given in Table 4. SVMs classifier had the best results.

Table 4. Performance Parameters of Classification.

<table>
<thead>
<tr>
<th>ML Algorithm</th>
<th>Accuracy (%)</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>AUC</th>
<th>F-Measure</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVMs</td>
<td>98.33</td>
<td>1</td>
<td>0.9667</td>
<td>0.9833</td>
<td>0.9818</td>
<td>0.9667</td>
</tr>
<tr>
<td>k-NN</td>
<td>96.67</td>
<td>1</td>
<td>0.9333</td>
<td>0.9667</td>
<td>0.9618</td>
<td>0.9333</td>
</tr>
<tr>
<td>Decision Trees</td>
<td>95.83</td>
<td>0.9833</td>
<td>0.9333</td>
<td>0.9583</td>
<td>0.9561</td>
<td>0.9167</td>
</tr>
</tbody>
</table>

3. RESULTS

A new algorithm was improved using only three features extracted from HRV for diagnosing HFrEF and HFpEF. At first, the number of extracted features was 37. But for real-time systems, extracting so many features was hard. So, to improve the performance, the amount of features was reduced. Feature ranking for all features was given in the table of Mann Whitney-U Test results, amount of features was diminished to three for improving the classification efficiency, and used. These three features used for classification were in the top three in the ranking.

As a result of the algorithm, high accuracy was acquired using the 10th, 20th, and 21st features. The classification results were given in Table 4. These results, which were acquired only using three features, can be argued as a significant system performance for use in practice.

We mentioned HFrEF-HFpEF binary classification in the results section of our triple classification study with PPG [23] before. However, in the study, lower accuracy was achieved with more features extracted from both PPG and HRV. Since this binary classification with PPG has not been studied before in the literature, we continued to study this subject. In this study, high accuracy classification was performed with 3 features extracted from only HRV.

Classifiers’ ROC curves were given in Figure 3. According to the figure, the sensitivities of SVMs and k-NN classifiers were very good, all classifiers were very close to ideal and gave good results.

The most generally made examination in cases with doubted HF is the echocardiogram. Since patients are mostly characterized with LVEF in many clinical studies, it is more important than being a manifestation of HF. LVEF is ordinarily identified by using the echocardiogram. Radionuclide angiography or radionuclide ventriculography (Multiple Gated Acquisition, MUGA) and Single Photon Emission Computerized Tomography (SPECT) are also employed seldomly. Those technical equipments are costly, each of them requires a specialist, and some of the techniques are invasive. Besides, conditions, in which reaching the devices is restricted, may occur. On the contrary, PPG is not pricy, easy to get, and responds quickly. So, the applicability of the HFrEF - HFpEF distinction using HRV was studied for alternating to expensive devices.
4. DISCUSSION

With this study, an algorithm for a system that supports the physician to diagnose, and didn’t require a specialist was intended to use only PPG for two HF subtypes which could be overlooked and difficult to diagnose. The algorithm was developed with PPG, that could be readily obtained also with smartphones, to be inserted into appliances in the future.

The results of our study showed that three features ejected from PPG-derived HRV might be utilized in the diagnosing of HFpEF and HFrEF and would ensure significant outcomes. The way to diagnose HFpEF and HFrEF with only one signal is opened with easily obtained PPG signals.

Before echocardiography, expensive tests, and laborious blood tests, required for the diagnosis of HFpEF and HFrEF, PPG usage will abate the workload.

In our study, by signal processing and machine learning methods, it was deduced that PPG-derived HRV might be utilized in the diagnosis of HFpEF and HFrEF. Compared to other diagnostic methods, PPG measurement is easier. So, this system is also advantageous for patient comfort.
In the literature, HFrEF and HFpEF classification studies using PPG or PPG-derived HRV have not been previously performed. It is thought that our study will contribute to the literature with this aspect. In one of HF-related studies conducted in the current year, low HRV was shown to be associated with an increased incidence of coronary heart disease (CHD) and HF in a cohort of postmenopausal women [25]. In the study, a risk analysis was performed for HF, not a diagnostic study, using HRV derived from 12-lead ECG and demographic information such as age, race, body mass, etc. and information about treatments administered. More features are used, which are obtained by more demanding methods in that study. So, the system we propose is more practical than that study.

The point that our study contributes to the literature is that HFpEF and HFrEF can be distinguished with a sole signal by using PPG. It is thought that it will pioneer to the next researches on the related subject. The used dataset wasn't very large. So, conventional machine learning methods were performed. Novel studies may be accomplished by applying more recent ML techniques with a bigger dataset.

ACKNOWLEDGEMENTS

The authors would like to thank Assist. Prof. İbrahim Kocayiğit and Assoc. Prof. Cahit Bilgin, from Sakarya University Faculty of Medicine, for their valuable contributions in data collection and labeling, and valuable referees who contributed to the development of the article with their evaluations. This study was supported by Sakarya University Scientific Research Projects Coordinatorship with the project numbered 2021-9-33-125.

REFERENCES


Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. European Heart Journal, 37, 2129–2200.


