Novel DERMA Fusion Technique for ECG Heartbeat Classification

Qurat-ul-ain Mastoi [1] , Teh Ying Wah [1], Mazin Abed Mohammed [2] , Uzair Iqbal [3] , Seifedine Kadry [4] , Arnab Majumdar [5] and Orawit Thinnukool [6],*

1   Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur 50603, Malaysia; quratulain.mastoi@siswa.um.edu.my (Q.-u.-a.M.); tehyw@um.edu.my (T.Y.W.)
2   College of Computer Science and Information Technology, University of Anbar, Ramadi 31001, Iraq; mazinalshujeary@uoanbar.edu.iq
3   Department of Computer Science, National University of Computer and Emerging Sciences, Islamabad Chiniot-Faisalabad, Chiniot 35400, Pakistan; uzair.iqbal@nu.edu.pk
4   Department of Applied Data Science, Noroff University College, NO-4608 Kristiansand, Norway; skadry@gmail.com
5   Faculty of Engineering, Imperial College, London SW7 2AZ, UK; a.majumdar@imperial.ac.uk
6   College of Arts, Media, and Technology, Chiang Mai University, Chiang Mai 50200, Thailand
*   Correspondence: orawit.t@cmu.ac.th

Abstract: An electrocardiogram (ECG) consists of five types of different waveforms or characteristics (P, QRS, and T) that represent electrical activity within the heart. Identification of time intervals and morphological appearance of the waves are the major measuring instruments to detect cardiac abnormality from ECG signals. The focus of this study is to classify five different types of heartbeats, including premature ventricular contraction (PVC), left bundle branch block (LBBB), right bundle branch block (RBBB), PACE, and atrial premature contraction (APC), to identify the exact condition of the heart. Prior to the classification, extensive experiments on feature extraction were performed to identify the specific events from ECG signals, such as P, QRS complex, and T waves. This study proposed the fusion technique, dual event-related moving average (DERMA) with the fractional Fourier-transform algorithm (FrlFT) to identify the abnormal and normal morphological events of the ECG signals. The purpose of the DERMA fusion technique is to analyze certain areas of interest in ECG peaks to identify the desired location, whereas FrlFT analyzes the ECG waveform using a time-frequency plane. Furthermore, detected highest and lowest components of the ECG signal such as peaks, the time interval between the peaks, and other necessary parameters were utilized to develop an automatic model. In the last stage of the experiment, two supervised learning models, namely support vector machine and K-nearest neighbor, were trained to classify the cardiac condition from ECG signals. Moreover, two types of datasets were used in this experiment, specifically MIT-BIH Arrhythmia with 48 subjects and the newly disclosed Shaoxing and Ningbo People's Hospital (SPNH) database, which contains over 10,000 patients. The performance of the experimental setup produced overwhelming results, which show around 99.99% accuracy, 99.96% sensitivity, and 99.9% specificity.

Keywords: cardiovascular disease; ECG signal processing; features extraction; machine learning; ECG heartbeat classification

1. Introduction

Cardiovascular diseases (CVDs) are the leading cause of death globally, as reported by the World Health Organization [1]. Sustainable development has been made in past years for reducing the impact of heart morbidity. An electrocardiogram (ECG) signal is the most widely utilized biosignal for CVDs detection on an early basis. The ECG is a non-invasive measurement of the heart that is utilized to diagnose different cardiac illnesses and anomalies [2,3]. Cardio specialists have been utilizing ECG waveforms for over seven

decades to distinguish heart illnesses, for example, arrhythmia and myocardial areas of dead tissue for more than 70 years [4].

Many cardiovascular diseases can be detected using the analysis of variation in ECG waves. However, the presence of artefacts such as baseline wandering and powerline interference in ECG signals can generate additional spikes in the waveform. Therefore, it is mandatory to eliminate the noisy elements from ECG signals to classify heartbeats accurately. The wavelet transforms technique is one of the most commonly used to remove noise and artifacts from ECG signals [5–7]. There have been several algorithms previously proposed to extract waveforms from the ECG cycle [8–11]. The authors in [12] proposed a rapid ramp effective algorithm to determine the occurrence of the R-peaks from ECG signals. This algorithm has higher-order complexity, and it is only applied to two records. Moreover, Hilbert transform and empirical mode decomposition are combined to detect the R-peaks from the ECG signal [13]. However, this approach is quite complex due to the consideration of massive amounts of blocks for detecting R-peaks. R peak detection is the sole purpose of both algorithms. In literature, some of the algorithms are proposed techniques to detect the P- and T-waves with R-peak [14,15]. These techniques include several steps, such as filtering the signals, creating interesting blocks for each peak individually and setting the fixed threshold point. To extract the large values from ECG signals, this algorithm specifically used the Butterworth filter to remove the baseline drift and square the output of the filtered signal. The selection of the area of the window depends on the length of the QRS complexes and repetition intervals. Each block's width is computed based on the size of the window and compared with the threshold. Eventually, peaks are detected from each width of the window. The values of the prior and after of the detected R-peaks are selected and samples of the R-peaks are adjusted to 0 for the identification of P- and T-waves. As far as peak detection is concerned, this algorithm yields acceptable results. In this study, we propose a novel dual event-related moving average-based fusion technique, which allows extracting the specific events from the ECG signals, including P-, QRS complex, and T-waves. However, these specific events of the ECG signal contain time intervals due to the time frequency-based analyses that produce huge variations related to data from the specific waves. Furthermore, this paper demonstrates that the proposed technique significantly achieves better performance than the state-of-the-art methods. Other major aim of this study is to classify five types of heartbeats (PVC, LBBB, RBBB, PACE, and APC) from ECG signals to identify the exact condition of the heart. This aim includes two stages: to extract the features and classification using the supervised model. MIT-BIH arrhythmia benchmark dataset has been widely used for many similar experiments in the last decade [16]. Previous research employed a variety of preprocessing approaches, feature extraction methods, and classifiers, some of which are covered in this study. In [17], discrete-wavelet-transform (DWT) was used to extract R-peak and RR interval, whereas multi-layer perceptron (MLP) was utilized for classification. The accuracy was achieved at 99.9%, with 301 features in classification. Similarly, in [18], db4 DWT was used to extract the R-peak location and RR interval, support vector machine classifier was used to detect the abnormality from ECG signals. In [19,20], various classifiers were utilized for the classification of ECG signals, such as naïve Bayes, SVM, ANN, and support vector machines.

The novel proposed technique has the potential to be applied to identify the normal and abnormal events of the ECG signals. This type of system involves implanted probe-less ECG sensors on the patient's body and transmitting the signals via Bluetooth connection to the processing device, for instance, a cellphone device, handheld device, and wearable wireless device. The obtained signal can be used to pre-process and extracts features for the classification of five heartbeat classes, namely PVC, LBBB, RBBB, PACE, and APC. The main highlights of the work are as follows:

1. We designed a fusion-based method DERMA with FrlFT to preprocess the ECG signal efficiently and extract the features such as P-, QRS, and T-waves.
2. Our proposed method specifically identifies the abnormal morphological events that are associated with PVC, LBBB, RBBB, PACE, and APC.

3. We calculated the temporal values, which are associated with specific events such as PR and RT waveform duration.
4. Furthermore, this study performed classification using supervised learning models, such as KNN and SVM, by utilizing the MIT-BIH Arrhythmia dataset, whereas two different datasets were utilized, namely AHA and SPNH, for testing purposes.

The organization of the paper is as follows, Section 2 discusses materials and methods, and Section 3 describes the classification of heartbeats. Section 4 discusses the results of the proposed model, and Section 4 provides the conclusion of this study.

2. Materials and Methods

To detect the accurate waveform from ECG signals, it is mandatory to preprocess the ECG signals as discussed in the introduction. Figure 1 represents the block diagram of the proposed methodology, which includes the following steps: (1) to eliminate the noise from the ECG signal, (2) to detect the specific events from the ECG signals using DERMA with the FrlFT fusion technique, and (3) to select the supervised model for the classification of ECG signals for the determination of CVD. Each task of the proposed methodology is briefly discussed in the following subsections.
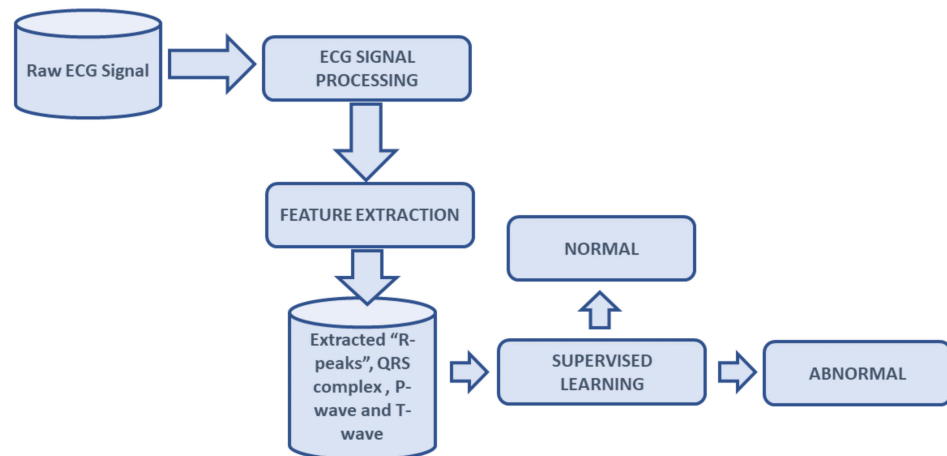

Figure 1. Block diagram of novel DERMA technique.

2.1. Preprocessing of ECG Signals

In this section, the first and foremost step is to normalize ECG signals to obtain the essential highest information. As we know, the presence of noise in ECG signals is non-linear and dependent on time whereas, amplitudes measurement is present in each subject, so it is necessary to calculate the statistical mean of the amplitude and then the difference among the $x(i)$ and μ by using the following equation:

$$x(i) - (\mu) \tag{1}$$

μ is the statistical mean of the amplitude; σ is the standard deviation of the sample. The Z-score method is used to normalize all amplitude values from the ECG signals

$$z = \frac{x(i) - (\mu)}{\sigma} \tag{2}$$

Discrete Wavelet Transform: determine the detailed co-efficient of ECG signals $x(i)$ DWT of a function that is defined as follows [21]:

$$W_\phi(S_o, k) = \frac{1}{\sqrt{M}} \sum_{k-0}^{M-1} x(t)\phi_{S_0,k}(t) \tag{3}$$

Additionally,

$$W_\psi(S,k) = \frac{1}{\sqrt{M}} \sum_{k-0}^{M-1} x(t)\psi_{S,k}(t) \tag{4}$$

It is stated that $\geq S_0$, where S_0 represents starting scale, and $\phi_{S,k}(t)$ and $\psi_{S,k}(t)$ are defined as the scaling and wavelet function. The detailed coefficient for the inverse discrete wavelet transform is presented below:

$$x(t) = \frac{1}{\sqrt{M}} \sum_{s_0=0}^{s-1} W_\phi(s_o,k)\phi_{s_0,k}(t) + \frac{1}{\sqrt{M}} \sum_{s=0}^{s-1} W_\psi(s_o,k)\psi_{s,k}(t) \tag{5}$$

Fractional Fourier Transform (FrlFT): The FrlFT is a generalized variant of the conventional Fourier transform that includes a parameter α. FrlFT is mostly employed in quantum physics to solve differential equations, but it may also be utilized to analyze optical-related problems [22]. In recent trends, advanced applications have proposed Fourier transform to preprocess the signal due to the useful characteristics in the time-frequency domain. The representation of the FrlFT signal is illustrated below [23]:

$$FrlFt^\phi(t,u) = F^\alpha(x(t)) = X_{]\phi}(u) = \int_{-\infty}^{\infty} x(t)K_\phi(t,u)dt \tag{6}$$

where α and $\phi = \alpha\pi/2$ represent the order of FrlFT and angle of rotation, respectively. The value of the integer is defined by n, while the FrlFT operator defined as $F^\infty(.)$, and kernel function of the FrlFT is computed by using the equation below:

$$k_\phi(t,u) = \begin{cases} \sqrt{\frac{1-jcot\phi}{2\pi}} \exp\left(S\frac{t^2+u^2}{2}cot]\phi - Stucsc\phi\right), \phi \neq n\pi \\ \delta(t-u), \ for \ \phi = 2n\pi \\ \delta(t+u), \ for \ \phi = 2\left(n+\frac{1}{2}\right)\pi \end{cases} \tag{7}$$

2.2. Signal Filtering

The nature of the ECG signals is non-stationary, which means that the frequency response of the ECG signals varies according to the time. Therefore, the contamination often varies according to the time-dependent variation. Thus, the conventional Fourier-transform technique does not provide the localization of the specific events of the ECG signal, whereas DWT extracts time-based features by removing noisy elements from ECG signals [18]. In order to remove baseline wander and power line interference efficiently from ECG signals, DWT uses two filters, namely low-pass and high-pass, for baseline wandering and power line interference, respectively. After applying these two filters, the central frequency component (CFC) is computed to identify the 0 to 1 wavelet ranges, according to the similarity ratio between the original signal and the wavelet-based chosen component. CFC is identified by using the Daubechies-4 (db4), which is around 0.7. Furthermore, the pseudo frequency (Fp) of each scale is computed by using the following equation:

$$F_p = \frac{F_c F_s}{2^p} \tag{8}$$

2.3. R-Peak Detection

The temporal localization might be lost by performing the Fourier transform on the ECG data. Thus, the purpose of implementing FrlFT is to identify the time-frequency-based features from the ECG signal [24]. FrlFT works as a rotational form when it rotates to the higher value of α, FrlFT extracts the frequency domain, and when it rotates towards the lower value of α, it extracts the time domain values of the ECG signals. To detect accurate high-frequency component, such as R-peak, time localization is the most important and necessary factor, which needs to be considered carefully [18]. In this study, we employed the

hit and trial technique for the value of α, and we found that it is equal to 0.0 l; appropriately, this process helps to enhance the R-peaks detection. The dual event-related moving averages technique helps to calculate the ECG cycle by using the following equation:

$$MA_{event}(n) = \frac{1}{W_1} \sum_{k=-l} x(n+k),$$ (9)

$$MA_{cycle}(n) = \frac{1}{W_2} \sum_{k=-p} x(n+p),$$ (10)

In this scenario, W_1 and W_2 represent the QRS complex and heartbeat durations, respectively. The optimum value of β factor was selected using the hit and trial technique, and it was used to multiply the mean (μ) of the enhanced signals. To generate the value of the threshold output number added to the MA_{cycle}. The value of the MA_{event} was compared with the associated threshold value. It is stated that if the value of $MA_{event}(n)$ was more than *nth* threshold element, 1 is assigned to the new vector; otherwise, it considers 0. The sequence of nonuniform heartbeat is constructed using this method. Moreover, that heartbeat, which has a width equivalent to the W_1, contains the block of the desired event. The maximum value of the enhanced signal is considered as R-peak in the interesting event block. The accurate detection of the R-peak using the proposed algorithm is shown in Figure 2. Our proposed model also extracts negative polarity of the R-peak, which is beneficial to detect different types of arrhythmias from ECG signals. Figures 3 and 4 represents the area of the interested block and two events results, respectively. To accurately identify the P- and T-peaks in the ECG signal, our proposed algorithm DERMA employs a simplified threshold value, which helps to minimize the computational complexity.
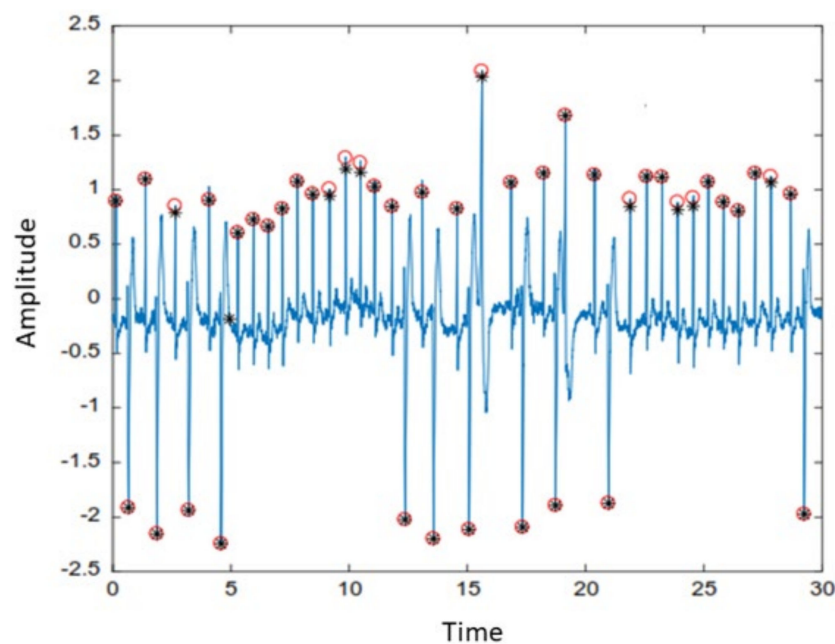

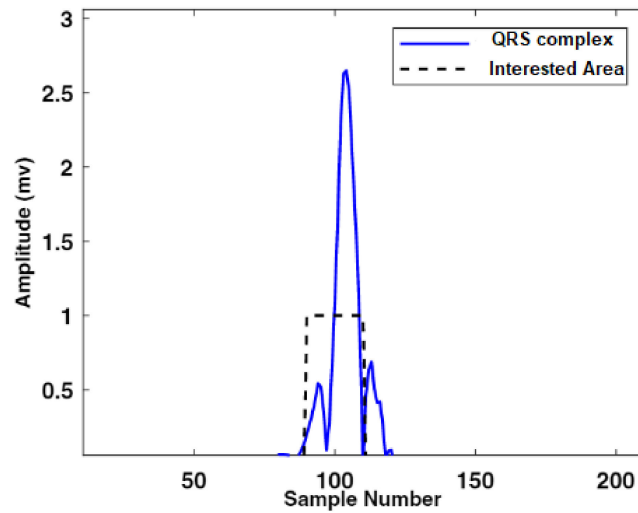**Figure 2.** The positive and negative polarity of R−peak detection.

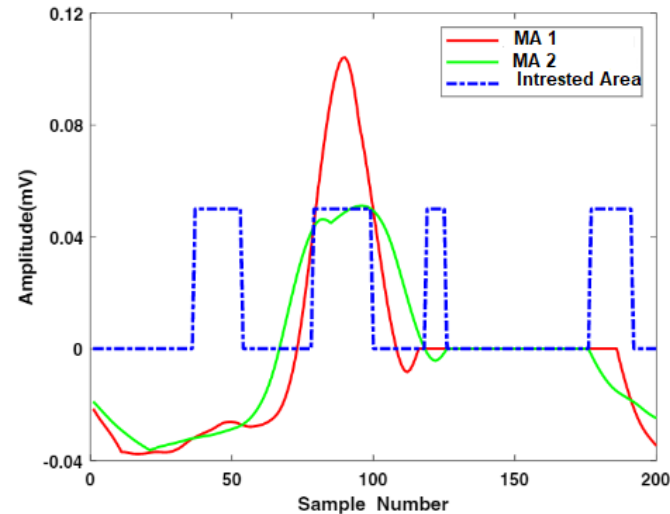**Figure 3.** Representation of QRS complex and area of interest.

Figure 4. Representation of QRS complex and area of interest using two events.

After the elimination process, a total of around 30 samples remained, which are 0.083 s before the eliminated part of the R peak, and 60 samples, which are around 0.166 s after the eliminated part of R; the peak was initialized to 0. However, P and T−peaks occur in the selected samples, and it was almost zero in all types of abnormal heartbeats. After the elimination of the QRS interested block, the FrlFT technique was used for prominent P and T-peaks for the time frequency-based feature extraction. In this scenario, W_3 and W_4 represent the peaks and duration of the waves, respectively. Figure 5 represents the extracted P and T-peaks. The following equation defines the DERMA for the calculation of the desired block.

$$MA_{peak}(n) = \frac{1}{W_3} \sum_{k=-q} x(n+q), \tag{11}$$

$$MA_{wave}(n) = \frac{1}{W_4} \sum_{k=-r} x(n+r), \tag{12}$$

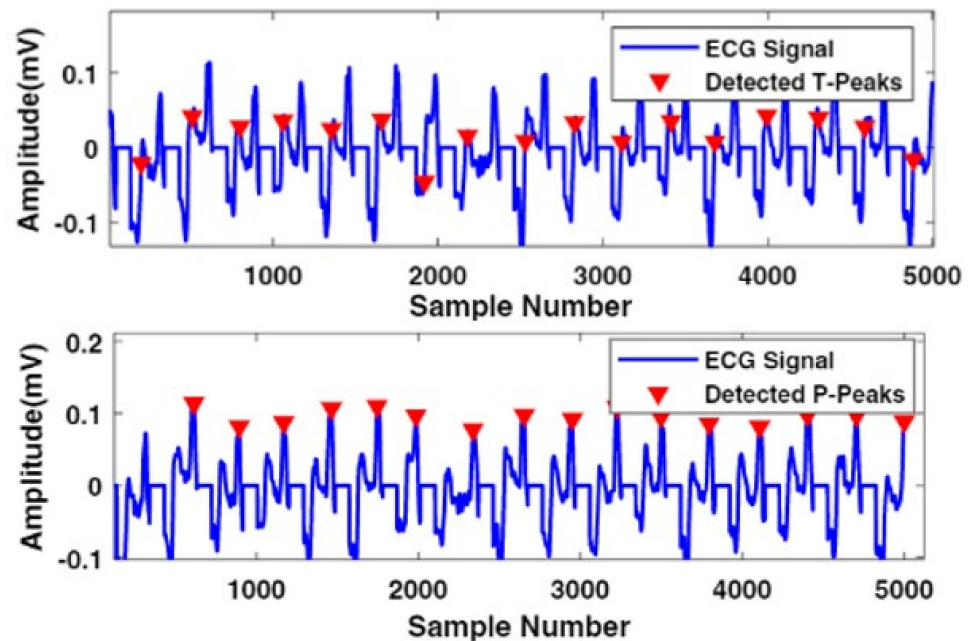$$q = \frac{W_3 - 1}{2} \text{ and } r = \frac{W_4 - 1}{2} \tag{13}$$

Figure 5. Graphical representation of P and T−peaks detection.

3. Classification

Machine learning was used in this section to classify the PVC, LBBB, RBBB, PACE, and APC heartbeats from the ECG signals. In general, machine learning algorithms are trained by using the labeled dataset, where various attributes from each dataset are extracted and formed into a model to predict the outcome. This method is called supervised learning. This method helps to automate the decision-making process by training and testing the different data. However, it includes two major processes, which are feature extraction and classification.

3.1. Feature Extraction

This is the unique way to recognize the pattern of diseases from ECG signals using the collection of different types of morphological changes in ECG signals, such as calculation of different peaks, finding intervals, or duration of peaks. These types of features or characteristics distinguish the normal and abnormal events in the ECG signals. Besides this, coefficients of the auto-regressive model help to identify the different characteristics of ECG signals [25]. The order of autoregressive model *p*, AR (*p*) is illustrated as:

$$x(n) = \sum_{i=1}^{p} a(i)x(n-i) + e(n) \tag{14}$$

In this scenario, $a(i)$ is the *i*th coefficient of the AR model, $e(n)$ is a white noise with a zero mean, and the order of AR model is represented as *p*. The selection of optimal order is determined by several parameters. The feature-extraction phase is most important because this step determines the better type of input in the ECG signal for classification.

Feature Matrix

This is the unique vector that contains the feature-related information of the ECG heartbeats. Two types of datasets are used in this experiment: MIT-BIH arrhythmia [16] and SPNH [26]. Each row contains different types of single ECG heartbeat-related features such as AR order, FrlFT, and interval duration of the peaks, such as QRS complex, P−wave, T−wave, and other temporal values, such as last R-R interval and subsequent R-R interval.

4. Supervised Machine Learning Algorithms

ECG signal classification is an important and difficult task. However, it provides detailed information about the abnormality in ECG signals automatically, which provides great benefit to society to reduce the impact of morbidity and mortality around the globe. This proposed model used an SVM and KNN supervised learning algorithm for this experiment. The following subsection briefly describes the machine learning algorithm implementation.

4.1. Support Vector Machine Classifier

SVM is under the category of supervised learning technique that mostly solves classification problems. This machine learning algorithm aims to provide the most powerful statistically based learning theory classifier. SVM facilitates the classification of unobservable patterns from the dataset [24]. It is a hyperplane that divides positive and negative sets with maximal margin. The dataset consists based on pairs $(a_1, y_1), (a_2, y_2), \ldots (a_N, y_N)$, where the hyperplane is demarcated as the following quadratic problem:

$$\max_{\alpha \geq 0} \left(\sum_{i=0}^{l} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{l} \alpha_i \alpha_j y_i y_j K(X_i, X_j) \right) \tag{15}$$

$$subject\ to\ \sum_{i=0}^{l} \alpha_i y_i = 0 \tag{16}$$

$$\alpha_i \leq C, i = 1, 2, \ldots, l, \tag{17}$$

In this equation X_i, X_j represent the input features, whereas y_i, y_j represent the class labels, and $\alpha_i \geq 0$ are Lagrangian multipliers, in which C is a constant variable, and K represents the kernel function [27].

4.2. K-Nearest Neighbor

KNN has used an instance-based learning method for storing the instances data for training the classifiers using similarity index measures. It is a versatile and robust classifier that is frequently used for more complex classification experiments. KNN essentially determines K instances in the training dataset by using the most popular distance measure technique, Euclidean distance. KNN is one of the simplest and most competitive algorithms, which achieves the highest performance in pattern recognition issues. KNN can give highly competitive results due to its simplistic nature [28].

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^{n} \left(a_r(x_i) - a_r(x_j) \right)^2} \tag{18}$$

Whereas the mathematical notation of KNN for Y is given as:

$$Y = \frac{1}{K} \sum_{x_i \in N_k(x)} y_i \tag{19}$$

where a_r and Y shows the value of rth variable of an instance x and local mean vector, respectively; $N_k(x)$ represent as the neighborhood of x element [29,30]. In the literature, the KNN classifier is mostly found for classifying abnormalities in the heart, or it is also used to detect heart diseases [20,31–33]. Moreover, this machine learning algorithm allows learning the patterns of the dataset from the training feature vector set by using the comparison of the similarity between the test and training feature vector set. Next, we compute the *k*-value using similarity measures that are used to determine the specific class from the dataset.

4.3. Evaluation Methods

The current research utilized a 10-fold cross-validation technique for the evaluation and validation of the model. This procedure is repeated until the k times final test result is

computed. Moreover, the performance of the DERMA fusion technique is assessed by the various common performance indicators, such as sensitivity (*Se*), specificity (*Sp*), accuracy (*Acc*), positive predictivity (*PPV*), and error rate (*er*). Four parameters of the confusion matrix are used to derive the performance indicator, namely correctly detected beats (true positive = *TP*), undetected beats (false negatives = *FN*), correctly undetected beats (true negatives = *TN*), and falsely detected beats (false positive = *FP*). These statistical indices are expressed as follows:

Sensitivity measures the average values of positive subjects that are correctly identified from the ventricular arrhythmia in the beat from ECG subjects. It is computed as:

$$Se = \frac{1}{N} \sum_{i=1}^{N} \frac{(TP)_i}{(TP + FN)_i} \tag{20}$$

Accuracy computes the rate of correctly classified ventricular arrhythmia classes out of the total number of ventricular heartbeat subjects. Accuracy is computed as:

$$Acc = \frac{1}{N} \sum_{i=1}^{N} \frac{(TP + TN)_i}{(TP + FP + TN + FN)_i} \tag{21}$$

Positive Predictive Value represents the average value of the subjects with positive ventricular arrhythmia in the heartbeat with truly or correctly classified ventricular arrhythmia. It is defined as:

$$PPV = \frac{1}{N} \sum_{i=1}^{N} \frac{(TP)_i}{(TP + FP)_i} \tag{22}$$

Error Rate is used to calculate the rate of ventricular arrhythmia in the heartbeat using the rate of incorrectly classified arrhythmias out of all the ventricular heartbeat details. It is represented as:

$$Der = \frac{1}{N} \sum_{i=1}^{N} \frac{(FP + FN)_i}{(TP + TN + FN + ssFP)_i} \tag{23}$$

5. Results and Discussion

The performance evaluation is the most important phase to validate model performance for ventricular heartbeat classification.

5.1. Simulation Results

In this section, we discuss the results of the two scenarios, including ECG feature extraction and classification of cardiovascular disease.

5.1.1. Feature Extraction Results

DERMA with FrlFT-based fusion technique detects various types of morphological points from ECG signals to efficiently diagnose the disease. The validation of the proposed method was performed using the MIT-BIH Arrhythmia dataset. The main advantage of the proposed algorithm is that it works independently for feature extraction, which means any lead can be selected for the extraction of high-frequency components from ECG signals. Table 1 shows the performance of both datasets, where a total of six types of features are extracted. The performance results here measured the successful detection of the features with true positive and true negatives, and the parameters in Table 1 identify the sensitivity and accuracy, which are observed from the successful detection of the features. Moreover, the detection error rate defines the error that DERMA fails to extract. DERMA fusion resulted in 99.99% sensitivity and accuracy in QRS complex detection and SDSD. Moreover, the extraction of P-wave showed 99.90% sensitivity and 99.99% accuracy. T-wave showed 99.92% sensitivity and 99.98% accuracy. The performance of detecting the previous RR interval and subsequent RR interval is observed at 99.99% and 99.97% sensitivity and

99.97% and 99.98% accuracy, respectively. Furthermore, average error detection is observed at around 0.0015%, which is quite less.

Table 1. DERMA Fusion performance results of feature extraction using both datasets.

S. No	Feature	Sen %	Acc %	Der %
1	QRS complex	99.99	99.99	0.002
2	P-wave	99.90	99.99	0.001
3	T-wave	99.92	99.98	0.002
4	The previous R-R interval	99.99	99.98	0.001
5	The subsequent R-R interval	99.97	99.98	0.001
6	The standard deviation of successive difference (SDSD)	99.99	99.99	0.002
	Total/Average Performance	99.96	99.98	0.0015

5.1.2. Classification Results

In this study, we trained the classifier using the MIT-BIH Arrhythmia dataset and tested the classifier using the SPNH dataset. Hence, it was observed that around 70% of the feature vector matrix was utilized for the training and 30% utilized for the testing purpose. A total of five types of heartbeat classes were selected to classify in this experiment, namely PVC, LBBB, RBBB, PACE, and APC. Moreover, SPNH contains four types of groups in the dataset, which include supraventricular tachycardia (GSVT), atrial fibrillation (AFIB), sinus bradycardia (SB), and sinus rhythm (SR). As we know, the machine learning algorithm performs better when we increase the quantity of data, and it was noticed that MIT-BIH Arrhythmia has only 48 ECG subjects. Therefore, we used the SPNH dataset, which contains 12 lead ECG signals from 10,646 patients at 10 s long. The main advantage of using this dataset is that it is already denoised, and the database contains 11 common rhythms and 67 different cardiovascular conditions. The parameter for the SVM classifier was set to $\gamma = \frac{1}{2\sigma^2}$ and adjusted to 2.44×10^{-4}. Hence, it was observed that the performance of the MIT-BIH arrhythmia dataset shows a better performance than the SPNH dataset, which is demonstrated in Table 2. Both datasets have different sampling frequency rates; due to maintaining integrity and simplicity, this study fixed the resampling rate to 128 Hz for both datasets. Therefore, each dataset has a similar type of noisy elements, such as baseline wander and power line interference. Moreover, Table 3 shows the performance comparison of the current study with state-of-the-art methods. According to the Table 3, it was observed that authors in [34,35] achieved almost similar results; however, the limitation of state-of-the-art methods is observed in that these studies only extracted QRS complex from ECG signals, whereas this study extracts all necessary specific events from ECG signals to identify the exact condition of the heart. Only the QRS complex could not provide extensive knowledge about the abnormality in ECG signals. Hence, it is a quite complicated task because we did not found any solid studies in this regard that utilized the SPNH dataset and MIT-BIH Arrhythmia dataset together. Therefore, to fairly compare the algorithm of this study, only the MIT-BIH Arrhythmia dataset was utilized because most of the studies used this dataset in the past for their experiment. In this study, we provide a unique solution to detect the exact condition of the heart by determining five classes of heartbeats, namely PVC, LBBB, RBBB, PACE, and APC. The computational complexity of the proposed model was identified by using three components, including DERMA with FrlFT, and classification models. The computational complexity of our proposed model was calculated around O ($N \, log_2 \, N$), where N represents the number of features. According to the perspective of the computational cost, we observed that our proposed DERMA with fusion technique provided more stability while changes in datasets occurred. Hence, it was also observed throughout the experiment that normalization plays a vital role in both dataset cases.

Table 2. Performance Results of the Classification Models.

Dataset	SVM			KNN		
MIT-BIH	**Acc (%)**	**Sen (%)**	**PPV (%)**	**Acc (%)**	**Sen (%)**	**PPV (%)**
Normal	0.92	0.98	0.98	0.94	0.97	0.92
PVC	1.00	1.00	1.00	0.98	0.95	0.97
PACE	0.95	0.97	0.99	0.99	0.94	0.96
AFIB	0.98	0.99	0.96	0.98	0.93	0.98
APC	0.99	0.98	0.97	0.99	0.96	0.99
SPNH						
GSVT	0.70	0.71	0.68	0.59	0.80	0.87
AFIB	0.83	0.82	0.87	0.85	0.88	0.89
SB	0.82	0.84	0.90	0.91	0.92	0.89
SR	0.78	0.79	0.75	0.74	0.79	0.78

Acc, accuracy; Sen, sensitivity; PPV, positive predictivity.

Table 3. Performance comparison using MIT-BIH Arrhythmia.

Studies	Acc (%)	Sen (%)	Sp (%)	PPV (%)
[34]	99.88	99.93	NA	99.95
[35]	NA	99.99	NA	99.97
[36]	NA	87	99	85
[37]	92.73	7.35	96.70	88.01
[38]	96.38	97.88	97.56	95.46
[39]	NA	78.8	NA	90.8
Proposed	99.98	99.96	99.9	99.98

6. Conclusions

DERMA with FrlFT-based fusion technique is thus proposed to extract the features from ECG signals. MIT-BIH Arrhythmia and SPNH datasets were used in this experiment. However, both signals have different sampling frequencies, so in this study, we resampled the ECG signal at 128 Hz to maintain the simplicity and integrity of the algorithm. To preprocess ECG signals, we used the traditional wavelet decomposition method to denoise signals, whereas the usage of the fusion model increased the performance of the detection of morphological values of ECG signals to classify the five classes of heartbeat efficiently. Six types of features are extracted, which identify the actual variation in ECG signals to simplify the detection of abnormal events. Moreover, this study utilized two supervised models for the classification, namely KNN and SVM, which performed overwhelmingly, and we observed that the performance using the MIT-BIH Arrhythmia dataset was more remarkable than the SPNH dataset. Both classifiers were trained and tested using the two datasets, which are quite challenging tasks to integrate; the purpose of using two datasets together was to enhance the size of the dataset to train the learning algorithm efficiently. On the other hand, the reason for using the SPNH dataset was that it contains already preprocessed dataset with unique features. It was observed that our preliminary results are promising and further improved the identification of the different types of abnormalities that can lead to mortality.

Author Contributions: Q.-u.-a.M., T.Y.W. and O.T., conceptualization, investigation, methodology, resources, software, and writing—original draft; Q.-u.-a.M., T.Y.W. and O.T, investigation, validation, review, and editing; M.A.M., U.I., S.K. and A.M., conceptualization, software, and writing—review and editing; Q.-u.-a.M. and O.T, writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Informed Consent Statement: Not applicable.

References

1. Subathra, M.S.P.; Mohammed, M.A.; Maashi, M.S.; Garcia-Zapirain, B.; Sairamya, N.J.; George, S.T. Detection of focal and non-focal electroencephalogram signals using fast walsh-hadamard transform and artificial neural network. *Sensors* **2020**, *20*, 4952.
2. Mastoi, Q.U.A.; Wah, T.Y.; Raj, R.G.; Iqbal, U. Automated Diagnosis of Coronary Artery Disease: A Review and Workflow. *Cardiol. Res. Pract.* **2018**, *2018*, 2016282. [CrossRef]
3. Alyasseri, Z.A.A.; Alomari, O.A.; Papa, J.P.; Al-Betar, M.A.; Abdulkareem, K.H.; Mohammed, M.A.; Kadry, S.; Thinnukool, O.; Khuwuthyakorn, P. EEG Channel Selection Based User Identification via Improved Flower Pollination Algorithm. *Sensors* **2022**, *22*, 2092. [CrossRef]
4. Abdi Alkareem Alyasseri, Z.; Alomari, O.A.; Al-Betar, M.A.; Awadallah, M.A.; Hameed Abdulkareem, K.; Abed Mohammed, M.; Kadry, S.; Rajinikanth, V.; Rho, S. EEG Channel Selection Using Multiobjective Cuckoo Search for Person Identification as Protection System in Healthcare Applications. *Comput. Intell. Neurosci.* **2022**, *2022*, 5974634. [CrossRef]
5. Zhu, B.; Ding, Y.; Hao, K. A novel automatic detection system for ECG arrhythmias using maximum margin clustering with immune evolutionary algorithm. *Comput. Math. Methods Med.* **2013**, *2013*, 658243. [CrossRef]
6. Mahesh, V.; Kandaswamy, A.; Vimal, C.; Sathish, B. Random Forest classifier-based ECG arrhythmia classification. *Int. J. Healthc. Inf. Syst. Inform.* **2010**, *5*, 1–10. [CrossRef]
7. Romero, V.B. ECG baseline wander removal and noise suppression analysis in an embedded platform. *Ene* **2010**, *16*, 38.
8. Elhoseny, M.; Mohammed, M.A.; Mostafa, S.A.; Abdulkareem, K.H.; Maashi, M.S.; Garcia-Zapirain, B.; Mutlag, A.A.; Maashi, M.S. A new multi-agent feature wrapper machine learning approach for heart disease diagnosis. *Comput. Mater. Continua* **2021**, *67*, 51–71. [CrossRef]
9. Rahman, A.U.; Saeed, M.; Mohammed, M.A.; Krishnamoorthy, S.; Kadry, S.; Eid, F. An Integrated Algorithmic MADM Approach for Heart Diseases' Diagnosis Based on Neutrosophic Hypersoft Set with Possibility Degree-Based Setting. *Life* **2022**, *12*, 729. [CrossRef]
10. Mastoi, Q.U.A.; Wah, T.Y.; Raj, R.G.; Lakhan, A. A novel cost-efficient framework for critical heartbeat task scheduling using the internet of medical things in a fog cloud system. *Sensors* **2020**, *20*, 441. [CrossRef]
11. Mutlag, A.A.; Ghani, M.K.A.; Mohammed, M.A.; Lakhan, A.; Mohd, O.; Abdulkareem, K.H.; Garcia-Zapirain, B. Multi-Agent Systems in Fog—Cloud Computing for Critical Healthcare Task Management Model (CHTM) Used for ECG Monitoring. *Sensors* **2021**, *21*, 6923. [CrossRef]
12. Celin, S.; Vasanth, K. Detection and classification of R-peak using Naïve bayes classifier. *Int. J. Eng. Technol.* **2018**, *7*, 397–403. [CrossRef]
13. Chen, L.; Hao, Y. Feature Extraction and Classification of EHG between Pregnancy and Labour Group Using Hilbert-Huang Transform and Extreme Learning Machine. *Comput. Math. Methods Med.* **2017**, *2017*, 7949507. [CrossRef] [PubMed]
14. Panigraphy, D.; Sahu, P.K. P and T wave detection and delineation of ECG signal using differential evolution (DE) optimization strategy. *Australas. Phys. Eng. Sci. Med.* **2018**, *41*, 225–241. [CrossRef] [PubMed]
15. Costa, R.; Winkert, T.; Manhães, A.; Teixeira, J.P. QRS peaks, P and T waves identification in ECG. *Procedia Comput. Sci.* **2021**, *181*, 957–964. [CrossRef]
16. Mark, G.; Moody, R. *MIT-BIH Arrhythmia Database Directory*; Massachusetts Institute of Technology: Cambridge, MA, USA, 1988.
17. Ayub, S.; Saini, J.P. ECG classification and abnormality detection using cascade forward neural network. *Int. J. Eng. Sci. Technol.* **2011**, *3*, 41–46. [CrossRef]
18. Mazidi, M.H.; Eshghi, M.; Raoufy, M.R. Detection of premature ventricular contraction (PVC) using linear and nonlinear techniques: An experimental study. *Cluster Comput.* **2020**, *23*, 759–774. [CrossRef]
19. Kumar, S.U.; Inbarani, H.H. Neighborhood rough set based ECG signal classification for diagnosis of cardiac diseases. *Soft Comput.* **2017**, *21*, 4721–4733. [CrossRef]
20. Padmavathi, K.; Ramakrishna, K.S. Classification of ECG signal during Atrial Fibrillation using Autoregressive modeling. *Procedia Comput. Sci.* **2015**, *46*, 53–59. [CrossRef]
21. Azariadi, D.; Tsoutsouras, V.; Xydis, S.; Soudris, D. ECG signal analysis and arrhythmia detection on IoT wearable medical devices. In Proceedings of the 2016 5th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 12–14 May 2016. [CrossRef]
22. Ozaktas, H.M.; Ankan, O.; Kutay, M.A.; Bozdagi, G. Digital computation of the fractional fourier transform. *IEEE Trans. Signal Process.* **1996**, *44*, 2141–2150. [CrossRef]
23. Sejdić, E.; Djurović, I.; Stanković, L. Fractional Fourier transform as a signal processing tool: An overview of recent developments. *Signal Processing* **2011**, *91*, 1351–1369. [CrossRef]

24. Pławiak, P.; Acharya, U.R. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. *Neural Comput. Appl.* **2020**, *32*, 11137–11161. [CrossRef]

25. Masetic, Z.; Subasi, A. Congestive heart failure detection using random forest classifier. *Comput. Methods Programs Biomed.* **2016**, *130*, 54–64. [CrossRef]

26. Zheng, J.; Zhang, J.; Danioko, S.; Yao, H.; Guo, H.; Rakovski, C. A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. *Sci. Data* **2020**, *7*, 48. [CrossRef] [PubMed]

27. Cuomo, S.; De Pietro, G.; Farina, R.; Galletti, A.; Sannino, G. A revised scheme for real time ECG Signal denoising based on recursive filtering. *Biomed. Signal Process. Control* **2016**, *27*, 134–144. [CrossRef]

28. Peterson, L. K-nearest neighbor. *Scholarpedia* **2009**, *4*, 1883. [CrossRef]

29. Fix, E.; Hodges, J.L., Jr. Discriminatory analysis—Nonparametric discrimination: Small sample performance. In *Project No. 21-49-004, Report No. 11, Contract No. AF 41 (129)-31*; USAF School of Aviation: Randolph Field, TX, USA, 1952.

30. Orozco-Duque, A.; Martinez-Tabares, F.J.; Gallego, J.; Rodriguez, C.A.; Mora, I.D.; Castellanos-Dominguez, G.; Bustamante, J. Classification of premature ventricular contraction based on Discrete Wavelet Transform for real time applications. In Proceedings of the Pan American Health Care Exchanges, PAHCE, Medellin, Colombia, 29 April–4 May 2013; pp. 1–5. [CrossRef]

31. Faziludeen, S.; Sankaran, P. ECG Beat Classification Using Evidential K-Nearest Neighbour. *Procedia Comput. Sci.* **2016**, *89*, 499–505. [CrossRef]

32. Kalaiselvi, C. Diagnosing of heart diseases using average k-nearest neighbor algorithm of data mining. In Proceedings of the 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 16–19 March 2016; pp. 3099–3103.

33. Gilani, M.; Eklund, J.M.; Makrehchi, M. Automated detection of atrial fibrillation episode using novel heart rate variability features. In Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, Orlando, FL, USA, 16–20 August 2016; pp. 3461–3464. [CrossRef]

34. Gupta, V.; Mittal, M.; Mittal, V.; Chaturvedi, Y. Detection of R-peaks using fractional Fourier transform and principal component analysis. *J. Ambient Intell. Humaniz. Comput.* **2021**, *13*, 961–972. [CrossRef]

35. Yaqoob, T.; Aziz, S.; Ahmed, S.; Amin, O.; Aluomini, M.S. Fractional Fourier Transform Based QRS Complex Detection in ECG Signal. In Proceedings of the ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; Volume 2020, pp. 931–935. [CrossRef]

36. Zhang, J.; Liu, A.; Liang, D.; Chen, X.; Gao, M. Interpatient ECG Heartbeat Classification with an Adversarial Convolutional Neural Network. *J. Healthc. Eng.* **2021**, *2021*, 9946596. [CrossRef] [PubMed]

37. Pucer, J.F.; Kukar, M. A topological approach to delineation and arrhythmic beats detection in unprocessed long-term ECG signals. *Comput. Methods Programs Biomed.* **2018**, *164*, 159–168. [CrossRef] [PubMed]

38. Xie, T.; Li, R.; Shen, S.; Zhang, X.; Zhou, B.; Wang, Z. Intelligent Analysis of Premature Ventricular Contraction Based on Features and Random Forest. *J. Healthc. Eng.* **2019**, *2019*, 5787582. [CrossRef] [PubMed]

39. Zhou, Z.; Zhai, X.; Tin, C. Fully Automatic Electrocardiogram Classification System based on Generative Adversarial Network with Auxiliary Classifier. *Expert Syst. Appl.* **2021**, *174*, 114809. [CrossRef]