

RESEARCH ARTICLE

TFDGiniXML: A Novel Explainable Machine Learning Framework for Early Detection of Cardiac Abnormalities Based on Nonlinear Time-Frequency Distribution Gini Index Features

MOHAMED AASHIQ¹, SHAIFUL JAHARI HASHIM^{1,2},
FAKHRUL ZAMAN ROKHANI¹, (Member, IEEE), MARSYITA HANAFI¹,
AND AHMED FAEQ HUSSEIN³, (Member, IEEE)

¹Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia

²Institute for Mathematical Research (INSPER), Universiti Putra Malaysia, Serdang 43400, Malaysia

³Department of Biomedical Engineering, Al-Nahrain University, Baghdad 10072, Iraq

Corresponding authors: Shaiful Jahari Hashim (sjh@upm.edu.my) and Mohamed Aashiq (aashiqmm@gmail.com)

This work was supported in part by the Geran Inisiatif Putra Siswazah (GP-IPS), Universiti Putra Malaysia (UPM) under Grant GP-IPS/2022/9735000; and in part by MyTIGER 2023 under Grant UPM.RMC.800-1/1/MATCHINGGRANT.

ABSTRACT Cardiovascular diseases (CVDs) are the leading cause of global death, with approximately 80% of such CVD mortalities occurring in low and middle-income regions. Early detection of cardiac abnormalities is essential for timely intervention and minimizing mortalities. Automated CVD detection methods are vital, particularly in areas with limited healthcare resources. However, most existing AI-based techniques lack three critical aspects: model interpretability, longer-duration analysis, and effective use of nonlinear time-frequency approaches, which are necessary for ECG signals due to their nonlinear, nonstationary, and multi-component nature. This study proposes explainable intelligent classifiers incorporated with a novel sequence of time-frequency energy Gini Index (GI) features from the QRS complexes of ECG signals to address these challenges and enable early-stage CVD detection. These features are extracted using the Choi-Williams Time-Frequency method, reporting the first instance application of GI measures to nonlinear time-frequency distribution (TFD) for ECG analysis. Features are computed from one-minute windows, covering 30 minutes of ECG recordings. These interpretable features provide clear insights into normal and abnormal ECG patterns. The proposed method was trained and validated using the MIT-BIH Arrhythmia and Fantasia-Normal databases. Eight machine learning classifiers, including SVM, Random Forest, XGBoost, Gaussian Naïve Bayes, KNN, LinearBoost, CatBoost, and Logistic Regression were tested. The best model achieved 100% sensitivity, 94.4% accuracy, 95.24% F1-score, 90% precision, and 92.59% AUC on the test dataset. High sensitivity ensures reliability for medical screening by reducing False Negatives, making the approach suitable for integration into any type of smart device for accurate online and offline monitoring of CVD abnormalities.

INDEX TERMS Cardiovascular disease (CVD), electrocardiogram (ECG), explainable AI (XAI), Gini index (GI), machine learning (ML), time-frequency distribution (TFD).

I. INTRODUCTION

The associate editor coordinating the review of this manuscript and approving it for publication was Li Zhang.

It is an undeniable fact that cardiovascular diseases (CVD) are becoming the main cause of mortality of humankind

across the globe. As reported by the World Health Organization (WHO), Every year CVDs have taken the lives of tens of millions of humans. CVDs if not discovered, have a prolonged latency period, and are challenging to treat once they develop [1]. Despite the considerable developments in CVD research as well as the advancements in diagnostic and treatment techniques, it remains the leading cause of mortalities for a prolonged period across the globe. It is estimated that about 80% of CVD-related mortalities happen in low and middle-income territories [2]. It is necessary to implement effective strategies for treating and intervening CVDs by detecting them at the earliest possible stages. There is a growing amount of evidence that indicates that extensive engagement in rehabilitation and recovery activities significantly enhances the prognostication of the victims. Some studies further demonstrate that involvement in such programs can reduce the risk of prolonged hospital readmission by close to 25% and the risk of mortality rate by more than 40% [3]. The electrocardiogram (ECG) is a vital pervasive and non-invasive tool used to observe cardiac conditions during the clinical phases as it provides useful information on cardiac health and pathology. This enables the specialists to take part in diagnostics in the early stages through real-time ECG monitoring. The time and the administrative processes taken for transferring between the General Practitioners (GP) to the Specialists will vary depending on the countries and regions based on the available health facilities. Not only are cardiologists involved in the diagnostics and medications of cardiovascular diseases, but neurologists and neurosurgeons also play a crucial role, particularly in managing stroke-related complications [4]. Hence, the initial screening process is crucial in guiding the patient toward the right direction for further relevant analysis and treatments.

Cardiac Arrhythmia is one of the major CVD types that can lead to other major complications. Arrhythmia is a condition of irregularities in the rhythms of the heart, where the heartbeats can be excessively fast (Tachycardia), too slow (Bradycardia), or irregular patterns. Certain types of arrhythmias pose a significant threat to life, while others may be chronic and irreversible [5]. If not intervene they might require expensive long-term healthcare management through medications and lifestyle changes. Several observations of clinical assessments have demonstrated that majority of the arrhythmias are either sudden or sporadic, lacking a consistent pattern in their timing and frequency. This unpredictable nature necessitates long-term continuous monitoring to detect arrhythmias as they occur and minimize the risk of underdiagnosis [6]. Given the variability and deviations in pathological ECG readings, the same underlying condition may present differently on ECGs across various patients or even within the same individual. Consequently, cardiologists need extensive expertise and considerable clinical experience to make accurate and reliable diagnoses. Identifying abnormal heart rhythms is a challenging task that demands significant effort from GPs during the initial medical screening process, and

fatigue can lead to missed or misdiagnosed cases, potentially delaying crucial treatment. To improve accuracy and efficiency in detecting heart rhythm abnormalities, there is a growing anticipation for automated technologies including Artificial Intelligence. Automated technologies have the potential to minimize the strains on doctors and enhance timely patient care.

These heart abnormal conditions are detected automatically using several techniques varying from Digital Signal Processing (DSP) techniques to modern-day machine and deep learning techniques, and the hybrid combination of both. Even though the prediction performances are higher in deep learning models, they lack in interpretability of diagnostic results and model generalizability, where it will perform poorly on unseen data. Further, Deep learning is more prone to model overfitting [7]. Most of the existing deep learning works have utilized a very small chunk duration of the ECG signals varying from a few seconds to a single minute. Analyzing longer-duration signals is crucial in accurately identifying abnormal conditions and reducing false negatives [8]. Some of the abnormalities with intermittent or transient occurrences may occur sporadically and require longer observations to detect [6]. Signal processing techniques vary from time domain (temporal) analysis, frequency domain (spectral) analysis, and Joint time-frequency (JTFD or TFD) analysis. Joint time-frequency methods are diversified into linear and non-linear approaches. Time domain analysis has severe limitations since it is not able to capture the frequency features. Similarly, frequency domain techniques would not be able to capture the time domain information. Hence, the usage of joint time-frequency techniques will enable the ability to capture both time and frequency (spectro-temporal) domain features simultaneously [9]. Joint time-frequency techniques are believed to be powerful tools in signal processing, especially in biomedical engineering, telecommunications, audio signal processing, and seismology. They perform better in analysing and understanding the non-stationary signals with dynamic frequency components such as ECG signals, audio signals, mechanical and vibration signals. TFD approaches can be further categorized into linear and non-linear approaches [9]. It is widely accepted that non-linear TFD techniques are more appropriate compared to linear techniques for analysing and representing ECG signals due to their non-stationary and non-linear nature. Furthermore, these techniques could help to overcome the noise and artifact effects in the signals. Hence, they perform better feature extraction from such signals [10]. Nonlinear TFD techniques like CWD more accurately calculate the energy contents of nonlinear signals compared to linear methods. This results in richer feature extraction, capturing transient and frequency-modulated components that might be lost in linear methods. Different classifiers handle non-linearity differently by using various approaches such as ensemble of trees, nonlinear kernels, nonlinear activation functions, and deep architectures. Advanced methods like

ensemble models can exploit these nonlinear structures more effectively. However, suitable models for different datasets can only be identified through rigorous experimentation, as these ML techniques are entirely data-driven, and their performance is highly dependent on the characteristics of the datasets. Nonetheless, many researchers still rely solely on either time-domain or frequency-domain features for automated cardiac anomaly detection due to the lack of familiarity and skills, limited computing resources, and the lack of awareness about the potential of TFD techniques.

Our study proposes a novel approach of using an existing nonlinear TFD technique incorporated with a sequence of novel Gini Index (GI) [11] feature vectors obtained from the nonlinear TFD energy concentration values of QRS complexes towards classifying the abnormal ECG signals from healthy signals. These novel nonlinear TFD energy Gini Index feature vectors have significantly increased the classification performance of some of the selected ML classifiers. The QRS wave segments of an ECG signal represent the rapid depolarisation of the ventricles, serving as the critical marker for analysing heart rhythm, detecting heart anomalies, and diagnosing cardiac conditions. This study further focuses on addressing the association between heart abnormalities and the variation measure of the energy content of the ECG signals and finally classifies the abnormal and normal cases. Choi-Williams Distribution (CWD), a well-known and well-performing TFD has been utilized to compute the energy concentration and the related variation measurements [12]. This would greatly help to interpret the final diagnosis results to the required parties. Moreover, a 30-minute duration of each ECG record was considered to incorporate a larger number of heartbeats to provide accurate, reliable, and consistent results. It has been suggested from several studies that analysing a long series of beats will improve the diagnosis of heart problems [13].

The principal contributions of this work could be listed as follows: 1. This work utilizes the Gini Index values of CW-TFD as feature vectors for an intelligent classifier to detect cardiac abnormalities during the initial ECG screening process. 2. The work enhances model explanation by analyzing Gini Index patterns that capture energy variations between normal and abnormal ECG signals. 3. It incorporates a larger number of heartbeats with 30-minute ECG records to ensure error-free screening and robust diagnosis. 4. The proposed method achieves a 100% sensitivity rate, a critical factor in medical applications for example screenings to eliminate false negatives. 5. Additionally, we provide a comparative analysis of classification performance across multiple machine learning models to validate its efficacy. The Gini Index is an exceptional sparsity index measure that is widely applied in machinery fault diagnosis and radar communication applications by adopting signal processing approaches. Nonetheless, it is already being adopted for time-domain features of ECG signals for cardiovascular classification tasks. As far as authors are aware, this is the first time, the Gini index is applied on nonlinear

TFD in ECG signal analysis and the subsequent CVD classification flow. Further, this study is among the tops to incorporate a larger number of heartbeats, approximately 183,000 for a single ECG lead using conventional machine learning-based CVD classification tasks. However, some of the recent deep learning-based CVD classification studies have utilized datasets containing millions of heartbeats but much on the shorter ECG duration unlike the 30-minute duration [14], [15], [16].

The subsequent sections of this manuscript are arranged sequentially as follows. First, Section II discusses the overview of the pertinent and related works and the foundational background of the CW-TFD. The comprehensive demonstration of the experimental setup and the corresponding methodologies are detailed under section III. Then, Section IV provides the findings of all experiments and discusses them in detail. Finally, the manuscript is concluded with derived conclusions and the recommended future directions under section V.

II. RELATED WORKS

A. SIGNAL PROCESSING AND AI TECHNIQUES IN ECG SIGNAL CLASSIFICATION

In recent times, the use of deep learning techniques in cardiac abnormality prediction has shown significant growth across the continents. The primary difference between conventional machine learning and deep learning algorithms lies in the approach to feature extraction. Machine learning relies on handcrafted text or numerical features as input, whereas deep learning models like CNN automatically extract features during the learning process [17]. Hence, several images or very large dimensions of matrices can be directly fed into the deep learning models without being constrained by the challenges associated with the multiple dimensions of the features. Nonetheless, the main challenge in deep learning models is the interpretability of the diagnosis results in medical settings. Recently, researchers have been adopting various techniques for creating Explainable AI models (XAI) to address the need of model interpretability in the medical and healthcare sectors [18], [19]. Moreover, a notable influencing factor is that many researchers employing deep learning models and AI techniques may lack background knowledge or expertise in the signal processing domain, as it is predominantly covered within a limited range of curricula.

Several deep learning models including single and multi-dimensional Convolutional Neural Networks (CNN) [14], [15], [16], [20], [21], [22], Recurrent Neural Networks (RNN) such as Long Short-Term Memory (LSTM) [23], [24] and Gated Recurrent Units (GRU) [25], hybrid combination of both CNN and RNN techniques [26], [27], [28], and more recently, Transformer based models [29], [30], [31] have been used in classifying major types of heart anomalies with better prediction accuracy values. However, the interpretability or the reasons behind the projected classes and the generalizability of the classification model for practical

applications on unseen data are prone to low performance. Moreover, the ECG signal durations are typically in a few seconds were considered by the researchers while employing deep learning models most of the time due to the limitations in computing resources and the time required for the training process. The use of time-frequency transformation or frequency transformation is limited while using deep learning models. They are solely used for noise removal and pre-processing or 1D to 2D conversion of input data which is required in 2D-CNN-based deep learning models. Linear-type time-frequency approaches such as wavelet transformation and short-time Fourier transform (STFT) are exclusively used for such objectives. The researchers have also indicated that this transformation helped them to enhance the classification performance of their models.

However, researchers persist in employing conventional machine learning algorithms in ECG signal classification applications instead of deep learning techniques for various reasons [32], [33], [34]. When signal processing techniques are meticulously implemented during feature engineering and extraction phases, they effectively reduce the feature set to lower dimensions that are more appropriate for machine learning algorithms as opposed to deep learning methods. Most of the researchers have been choosing to employ deep learning techniques primarily to circumvent the need for handcrafted feature extraction. Deep learning will accomplish this autonomously for them with minimal human intervention. However, handcrafted features would significantly aid in the interpretation of diagnostic results, particularly in medical settings. Furthermore, the computational overhead and training duration are minimal when implementing conventional machine learning models.

Time-domain analysis, frequency-domain analysis, and time-frequency domain analysis are the main categories of signal processing techniques. They have been widely used in ECG signals, Electroencephalogram (EEG) signals, telecommunication signals, audio signals, seismology signals, and vibration signals for numerous years. In [35], researchers have extracted time domain features from QRS wave segments and ST-T segments of the ECG signals and used machine learning classifiers such as Support Vector Machines (SVM), Decision Tree (DT), and K-means clustering to cluster the Ischemia, Arrhythmia, and Healthy datasets. They used the venerable but limited Pan-Tompkins time-domain algorithm for detecting the R-peaks of the ECG signals. The same group of researchers in their subsequent work [36] have used both time domain features and frequency domain features for classifying the same classes of abnormalities. They have utilized Naïve Bayes (NB) and Linear Discriminant Analysis (LDA) classifiers by using these extracted features subsequently. In [8], researchers have used the CWD method to extract time-frequency features from PR and ST segments of ECG signals to detect the Ischemia cases and

healthy cases from several databases. They used an SVM classifier for classifying the cases. R.R. Sharma et al. used the TFD approach using a combination of Hilbert Transform and Hankel Matrix to extract multiple local and global features [37]. These features are derived from Average time frequency, frequency average, and time average from the TFD matrix. Thereafter, all features were inputted into the Random Forest and Decision Tree to classify the healthy and CVD patients. Zaid Abdu et al. proposed a heart sound classification method using photoplethysmogram (PPG) signals for identifying abnormal heart sounds [38]. Initially, they used fractional Fourier transform, a linear TFD approach to obtain the Mel-frequency coefficients and then the obtained features were fed into multiple classifiers such as KNN, SVM, and ensemble classifiers for the final classification. Researchers have also used the Fractional Fourier Transform to obtain multiple time-frequency domain features from the ECG signals and subsequently fed them to the SVM and Multilayer Perceptron (MLP) classifiers to categorize multiple types of Arrhythmia abnormalities [39]. They have also utilized Discrete Wavelet Transform (DWT) for noise and artifact removal before applying Fractional Fourier Transform. Another joint time-frequency approach was investigated in the experiment of extracting fetal ECG signals from abdominal ECG signals of pregnant mothers [40]. They further validated their method with the Arrhythmia dataset as well.

They have applied the Stockwell transform (S-transform) to retrieve the Shannon energy entropy features. The use of nonlinear time-frequency techniques is relatively minimal when compared to linear TFD and time-domain techniques in cardiac anomaly detection applications. Nonlinear approaches are more crucial in analyzing signals like ECG that are non-linear and non-stationary. Hence, in this work, we are proposing a nonlinear time-frequency approach for extracting the features and subsequently, machine learning algorithms for the final ECG anomaly classification.

B. CHOI-WILLIAMS TIME-FREQUENCY DISTRIBUTION (CW-TFD)

The underlying objective of using a time-frequency distribution technique is to derive a function that simultaneously represents the energy density of a signal in both time and frequency domains. This approach is valuable for many applications involving signals with time-varying spectra [13]. Joint TFDs are especially important for analyzing nonstationary and multicomponent ECG signals [40]. Since there is no single, universal way to depict the energy distribution of a signal across time and frequency, various methods have been proposed [12]. One of the most widely recognized methods is the quadratic, or Cohen's class, representation. Nonetheless, this approach often suffers from interference between frequency components, known as cross-terms. To minimize these unwanted cross-terms, one of the most effective

techniques is the Choi-Williams class kernel (CWD) for time-frequency analysis [12]. This can be expressed using (1).

$$CWD(t, \omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} K(\theta, q) x\left(m + \frac{q}{2}\right) x^*\left(m - \frac{q}{2}\right) e^{-j(\theta t + \omega q - s\theta)} dm dq d\theta \quad (1)$$

where, t and ω symbolize frequency and time components respectively. $K(\theta, q)$ is the exponential kernel function utilized in the CW-TFD. The kernel function could be considered as a low-pass filter, suppressing the cross-term interference effectively [8]. This kernel function of CWD is defined by (2) as below.

$$K(\theta, q) = \exp\left[-\frac{\theta^2 q^2}{\sigma}\right] \quad (2)$$

Here, σ is a real value parameter that can adjust the resolution and minimize the cross-term effects. The discrete formulation version of CWD for the sampled ECG signal data $x[m]$ could be defined mathematically as below in (3) [8],

$$CWD(n, \omega) = 2 \sum_{q=-\infty}^{\infty} e^{-j2\omega q} \sum_{m=-\infty}^{\infty} \frac{\sigma}{4\pi q^2} e^{-\frac{\sigma(m-n)^2}{4q^2}} x(m+q)x^*(m-q) \quad (3)$$

Another technique was suggested in [41] to further reduce cross-terms in Cohen's time-frequency classes by ensuring that only the positive components of the matrix retain the underlying information, whereas the negative components, which correspond to cross-term effects, are excluded from the CWD matrix output. And this approach successfully applied in various fields including in the analysis of ultrasound and electromyogram (EMG) signals for much-reduced computing and memory reasons for calculating CWD. The signals exhibit frequency components that vary over time, and the power density or marginal conditions are represented by the instantaneous spectrum [42]. The instantaneous power can be obtained by integrating the TFD over the frequency components as below in (4).

$$\int_{-\infty}^{+\infty} TFD(t, \omega) d\omega = |x(t)|^2 \quad (4)$$

Similarly, the integration of TFD function over the range of time components yields the energy spectrum of this TFD as follows in (5):

$$\int_{-\infty}^{+\infty} TFD(t, \omega) dt = |X(\omega)|^2 \quad (5)$$

Hence, the total power or energy can be derived as follows:

$$\begin{aligned} Energy_{(Total)} &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} TFD(t, \omega) d\omega dt \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} CWD(t, \omega) d\omega dt \end{aligned} \quad (6)$$

various distributions from Cohen's classes are used to analyze bio signals, seismic signals, mechanical vibration signals, etc. The CWD provides dependable and precise outcomes,

striking an effective balance between minimizing unnecessary cross-term effects and maintaining the strength of the TFD [9].

III. EXPERIMENTAL METHODOLOGY

A. ECG DATABASES

There are various ECG databases available to the public for ECG-based research works. They are mainly from the PhysioNet public repository [43]. These databases are considered golden standard databases for such tasks. We have selected two databases that were more than 30 minutes in duration due to the reason that all experiments were performed up to 30 minutes for each record by dividing them into one-minute windows. MIT-BIH Arrhythmia [44] and Fantasia Normal databases [45] from the Physionet public repository were used. The MIT-BIH Arrhythmia database comprises 48 Arrhythmia records from 47 patients for 30-minute durations. 25 male volunteers (Ages between 32 and 89) and 22 female volunteers (ages between 23 and 89). This took around 4 years for the cardiologists to process and annotate the beats.

Fantasia Normal includes the records of 40 healthy people with 20 young adults aged between 21 and 34, and 20 elderly subjects aged between 68 and 85. They were observed for 120 minutes while lying in a supine position, during which continuous electrocardiogram (ECG) and respiration data were recorded. Uncalibrated continuous non-invasive blood pressure measurements were also taken for half of each age group. Both subgroups consist of an equal amount of male and female subjects. MIT-BIH Arrhythmia records are considered for abnormal cases, and the Fantasia Normal records are considered as normal or healthy cases. Class labelling is accomplished accordingly. Thirty minutes of the MIT-BIH Arrhythmia database accounts for 109,494 beats for a single lead and the first 30 minutes of the Fantasia Normal database accounts for 73,510 beats approximately. The total duration of the Fantasia database lasts for 2 hours for each record. MIT-BIH Arrhythmia database includes two ECG leads, lead ML II and V5. However, the Fantasia Normal database contains only a single ECG lead, and the other two signals are the respiratory signal and the blood pressure waveform.

B. ECG SIGNAL PROCESSING

The collected ECG datasets were to undergo meticulous preprocessing prior to extracting the joint-time frequency features of QRS complexes. A resampling step was performed ahead due to the differences in the sampling rates across the selected databases. All ECG data were resampled at 360 samples per second as it is one of the typical sampling rates in related ECG studies. American Heart Association (AHA) recommends selecting the appropriate sampling rate for the accurate measurement of R-peaks and the range of 250-500 Hz or perhaps even higher rates are optimum [46], [47]. However, this would affect time-domain-based signal

analysis, on the contrary, it will not have any impact on our time-frequency approach as we are not performing any types of R-peak detection in the time domain. However, we maintain a higher sampling rate as they are more informative for discrete or digital analysis. The digital noise filters were applied to get rid of ECG artifacts and enhance the signal quality. A first-order high-pass Infinite Impulse Response (IIR) filter was applied to eliminate the baseline wandering effects, while a bandpass filter, with bandwidth at 0.5 Hz and 40 Hz of center frequencies, was employed to suppress the muscle noise artifacts and the high-frequency powerline interferences. The ECG raw data were also scaled to mitigate dynamic fluctuations caused by physiological differences and activities. This scaling step was crucial for ensuring uniform comparisons across different databases and for consistent feature extraction in a generalized algorithm.

Algorithm 1 illustrates the necessary steps for the elimination process of baseline drift and suppression of high-frequency artifacts.

Algorithm 1 Algorithm for Eliminating Baseline-Wander and High-Frequency Noise

Input: input_signal, sampling_rate, filter_coeff, cutoff_bp_low, cutoff_bp_high, filter_order
Output: final_filtered_signal

- 1: Initialize: Design the IIR high pass filter with filter coefficients filter_coeff
- 2: filtered_signal_hp $\leftarrow$ Apply IIR Filter (filter_coeff, input_signal)
- 3: bp_coefficients $\leftarrow$ Design Bandpass_Filter (cutoff_bp_low, cutoff_bp_high, sampling_rate, filter_order)
- 4: final_filtered_signal $\leftarrow$ Apply Bandpass_Filter (filtered_signal_hp, bp_coefficients)
- 5: **return** final_filtered_signal

Figure 1 visualizes the ECG signal before and after performing the noise removal steps. It can be noticed clearly from the figure that after removing the baseline wanders, the isoelectric level of the ECG signal is distinctly defined in contrast to the baseline wander-affected signal. This is crucial for precise measurements of the parameters and to avoid misleading clinical interpretations. The spectral contents of baseline wanders are usually confined to a frequency range below 0.5 Hz [48]. It is also visible from the figure that high frequency which is more than 40 Hz artifacts are removed after performing the bandpass filtering.

C. MEMORY AND COMPUTATIONAL REQUIREMENTS FOR QUADRATIC TIME-FREQUENCY DISTRIBUTIONS

TFDs are represented as two-dimensional functions, which demand significant computing and memory resources for analyzing larger datasets. Typical quadratic TFD algorithms

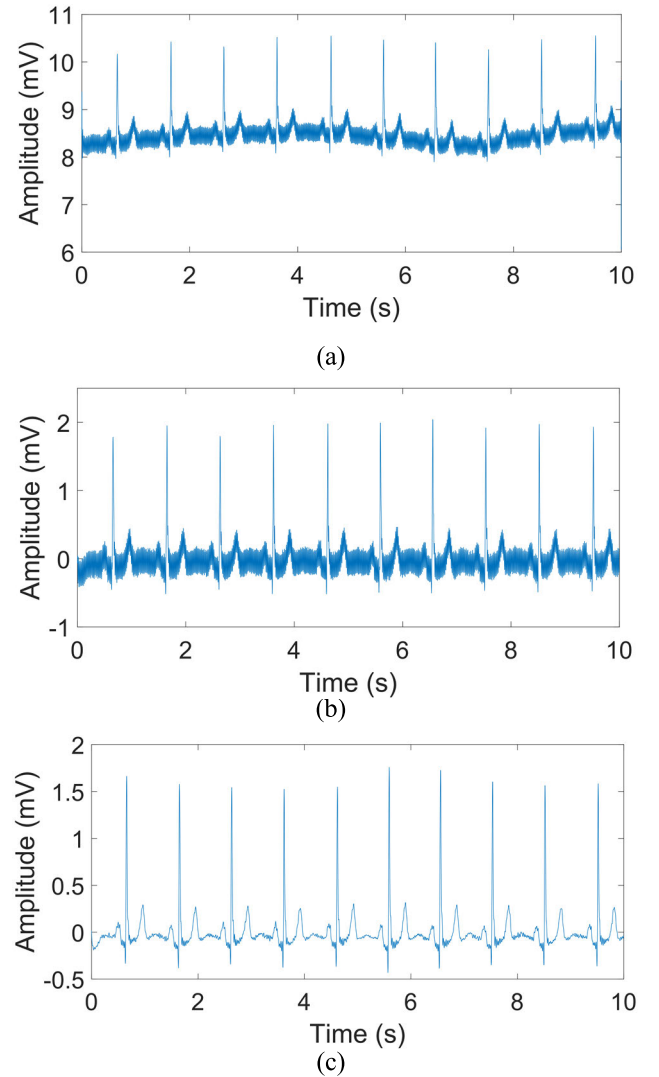


FIGURE 1. (a) ECG signal affected by baseline wander and high-frequency noises. (b) ECG signal after eliminating the baseline wanders. (c) Cleaned ECG signal after eliminating both baseline wanders and high-frequency noises.

hold approximately $N^2 \log_2 N$ operations, and $2N \times N$ size of memory. The real signal $S(t)$ is converted into an analytical signal $Z(t)$ before forming the TFD function. This will eliminate the negative frequency components and prevent cross-term interference between negative and positive frequencies in the time-frequency domain. The general mathematical form of a quadratic class distribution for an analytical signal $Z(t)$ can be expressed from the Equation (7) below.

$$\rho_Z(t, f) = \mathcal{D}_Z(t, f) *_t *_f \gamma(t, f) \quad (7)$$

where, $*_t *_f$ represents TF convolution, $\mathcal{D}_Z(t, f)$ represents the TFD, and $\gamma(t, f)$ is the TF kernel function. In this study, the approach from [49] that reduces memory and computational load is applied. This algorithm decreases memory usage from $2N \times N$ times to $N \times N$ times sample

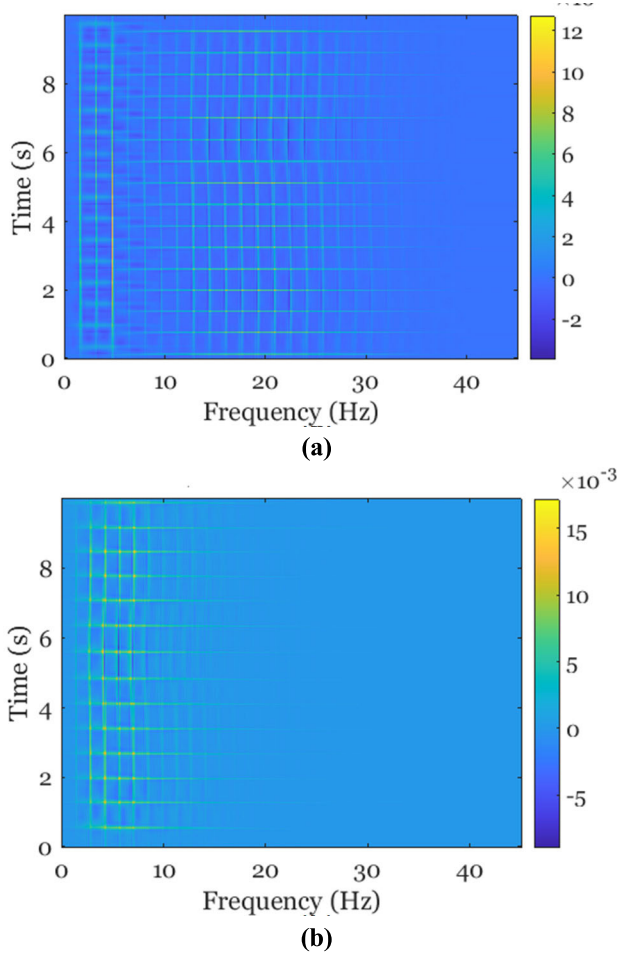


FIGURE 2. Two-dimensional (2D) representation of Choi-Williams transformation for (a) normal ECG and (b) abnormal ECG signals.

points and the computation load is reduced from $N^2 \log_2 N$ to $1/2N^2 \log_2 N$ numerical calculations. Despite using a memory-efficient algorithm, the need for additional memory remains, especially when analysing long-duration data. For instance, one minute of data sampled at 360 samples per second requires $(60 \times 360)^2 \times 8$ bytes, which equals approximately 3.47 GB of computer memory. Figure 2 shows the differences between healthy and abnormal ECG signals using CWD in 2-D representation. Similarly, Figure 3 visualizes the same in 3-D representation. It can be noted from Figure 3 that the energy distribution of QRS complexes in healthy subjects is well-defined and localized in both time and frequency domains. The 3D plot exhibits smooth and sharp peaks corresponding to the periodic occurrence of QRS complexes, indicating stable cardiac activities. Least background noises or spurious energy concentrations reflect efficient cardiac electrical conduction and regular heartbeats. In contrast, abnormal cases show broader and more scattered energy distributions. The QRS energy peaks appear more dispersed in both time and frequency, indicating irregularities in electrical conduction. Additional energy fluctuations outside the QRS region suggest

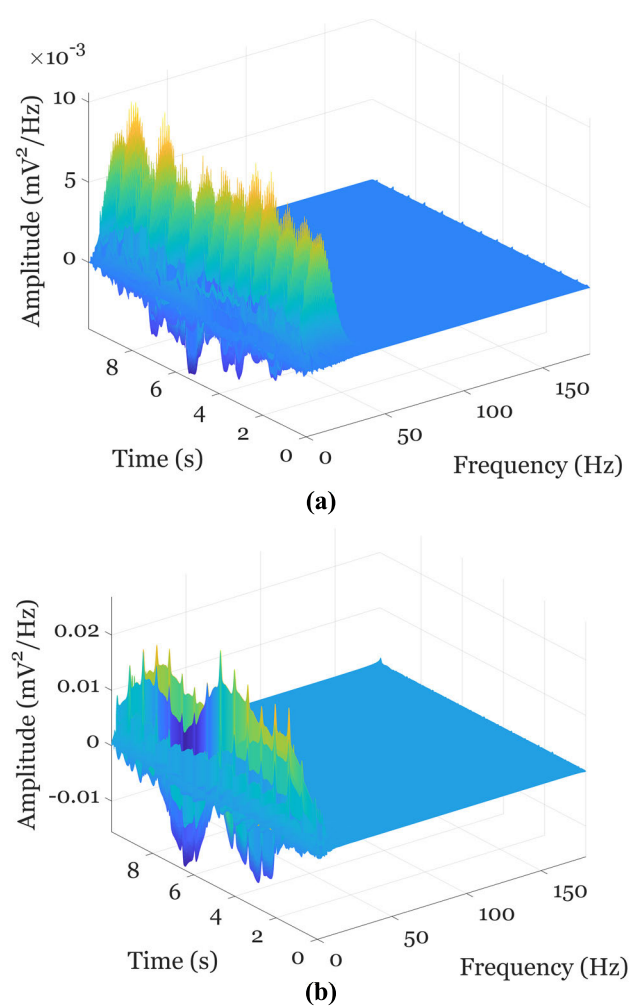


FIGURE 3. Three-dimensional (3D) representation of Choi-Williams transformation for (a) normal ECG and (b) abnormal ECG signals.

commonly associated cardiac abnormality variations. Some secondary energy components are linked to repolarization abnormalities.

D. EXPERIMENTAL SETUP

All experiments were performed in a high-end workstation equipped with 128 GB of RAM and 12 Physical CPU Cores (24 Logical Processors). The operating system used was Windows 10. An NVIDIA GeForce RTX 2080 Ti with 74.8 GB of GPU memory was also included in the same PC. All signal processing tasks were performed in MATLAB R2023a version, and the Machine Learning classification was performed in Python version 3.11.7 and the scikit-learn version 1.2.2.

E. TIME-FREQUENCY ENERGY CONCENTRATIONS

The QRS segments of the ECG signals, which lie within the frequency range of 5 to 22.5 Hz are the primary region used to ascertain the normality or abnormality. Total energy concentration will be computed by setting up the time slots and frequency slots over the TFD function obtained from the

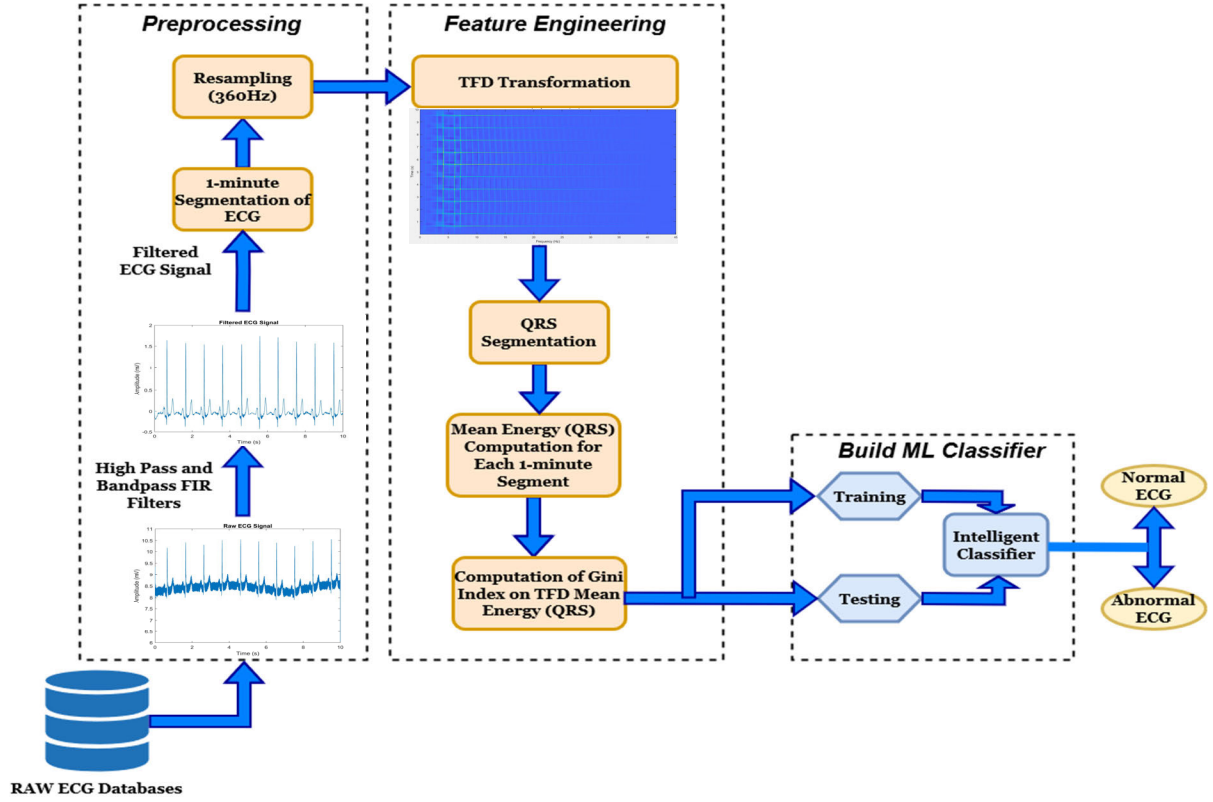


FIGURE 4. Methodology flow diagram for the proposed CVD anomaly detection system.

above CWD function given in Eq. 1 and computed as shown below in (8).

$$TE_{QRS}(t, f) = \left(\sum_{t=1}^N \sum_{f=5}^{22.5} CWD(t, f) \right) \quad (8)$$

Similarly, the marginal energy at the instant frequency f_i is computed as per the below in (9),

$$E_{QRS}(t, f_0) = \left(\sum_{t=1}^N CWD(t, f_0) \right), f_0 \in [5, 22.5] \quad (9)$$

Here, N indicates the total number of samples for the one-minute window of the ECG signal.

F. FEATURE EXTRACTION FOR TIME-FREQUENCY ENERGY GINI COEFFICIENTS

Feature extraction and feature engineering are essential in machine learning as they involve forming new and domain-specific features or modifying existing ones for a better representation of the underlying patterns in the data. Researchers have used multiple instances of a single feature [8], [9], [50], [51], [52] or single instance of multiple features or multiple instances of multiple features [5], [53], [54] in their respective works and achieved commendable classification performances in all these three approaches. In this work, we are proposing a sequence of Gini index

values or Gini coefficients as the feature vector to be used with several machine learning models.

These Gini index measures are obtained from the energy concentrations of QRS complexes by using the CW-TFD function as described in the above sections. Gini index measures can be used to indicate the fluctuations in QRS complex energy concentrations of TFD representation of abnormal and normal ECG signals. After computing marginal time-frequency energy values of all QRS complexes, the Gini index measures are computed for each one-minute window throughout the whole 30-minute duration for each record. Hence, 30 instances of GI measures are extracted and used as the input feature vector for the specified machine-learning classifiers.

The Gini index is commonly regarded as an exceptional sparsity index measure by researchers. It was introduced by Conrado Gini, an Italian Mathematician to study the income variation of a given society, and it originated from the Lorenz curve. GI exhibits the most consistent gradient behavior and outperforms commonly used sparsity statistical measures including Kurtosis, Lp/Lq norm, and Hoyer measures [55]. The Gini index has been widely used in economics and sociology to study mainly the wealth inequalities among populations. It is also well utilized in other areas of applications, including but not limited to detecting faults in machinery [55], gene studies [56], bacterial studies [57], healthcare resource distribution [58], pathological

studies such as tumor detection [59] and cardiovascular studies [60]. However, GI measure has been applied only in the time domain features of the ECG signals such as R-R intervals and other heart rate variability (HRV) measures [60]. To the best of the authors' understanding, this is the first occurrence of the GI measure being applied to study the variations of QRS segment energy concentration derived from a nonlinear joint TFD for ECG signal analysis.

The respective GI expression could be incorporated with the sum of energy concentration values as indicated below,

$$GI(E_{f0}) = 1 - \frac{2}{|E_{f0}|_1} \sum_{i=1}^N E_{f0s}(i) \left(\frac{N-i+1/2}{N} \right) \quad (10)$$

where, $E_{f0} = E_{QRS}(t, f_0)$ and $|E_{f0}|_1$ is the L1 norm of the E_T Values. $E_{f0s} : |E_{f0s}(1)| \leq \dots \leq |E_{f0s}(N)|$ are the ascending order sorted values of elements E_{f0} . The GI values range between 0 and 1 ($GI \in [0, 1]$), Where 0 represents the equal distribution and 1 signifies the complete inequality in the distribution. Hence, this GI measure would be more appropriate for concentration measurements compared to other commonly used metrics. The GI is useful for identifying the representation of most concentrated TFD, particularly, for the signals that exhibit intrinsic and nonlinear characteristics in the time-frequency domain. Furthermore, it could serve as an effective metric for determining the optimal parameter values of the TFD that best align with the analyzed signal. After obtaining the GI values of all subjects, they will be fed to the final intelligent classifier to determine whether the subject is healthy or has any kind of heart abnormalities. Algorithm 2 describes the complete sequence of computation of Choi-Williams transformation, Marginal energy concentration of QRS segments, and the extraction of Gini Index feature vectors of the TFD distribution. Figure 4 illustrates the comprehensive workflow of the proposed normal-abnormal ECG classification problem, which is described in the above Sections B, E, and F. Initially, the collected signals underwent high pass and bandpass filtering for noise and artifacts removal, following that, signals were segmented into one-minute window frames and resampled to 360Hz rates. After that, these ECG signals were transformed into CWD-TFD form and the mean energy values of QRS components were computed subsequently. From the computed mean-energy values, Gini-Index measures were derived respectively. These Gini-Index values are used as the feature vector for training and testing the provided intelligent classifiers subsequently. Finally, normal and abnormal ECG classification is accomplished by using these GI feature vectors.

G. INTELLIGENT CLASSIFIERS

Eight supervised machine learning algorithms are used as the intelligent classifier and their classification performances were compared with each other. All Gini Index feature vectors are labelled either abnormal or normal according to

Algorithm 2 Algorithm for Computing Choi-Williams Transformation and Extracting Gini Index Values from Temporal-Spectral Energy Concentrations

Input:

ecg_input: ECG signal data in Array format
 f_0 : The original sampling rate of the ECG signal
 f_t : The target sampling rate for resampling
 qrs_range: Frequency range of the QRS complex
 W : Duration of each window segment
 N : Number of 1-minute segments

Output: gini_coefficients

```

1: resampled_ecg_signal ← Resample(ecg_input,  $f_0$ ,  $f_t$ )
2: for  $i = 1$  to  $N$  do
3:   seg_start =  $(i - 1) * W * f_t$ 
4:   seg_end =  $i * W * f_t$ 
5:   ecg_segment ← resampled_ecg_signal
                     (seg_start: seg_end)
6:   cwd ← Compute TFD(ecg_seg) using the
                     given Equation (1).
7:   qrs_freq_range = [ $lf\_qrs$ ,  $hf\_qrs$ ]
8:   for  $j = lf\_qrs$  to  $hf\_qrs$  do
9:     marginal_qrs_energy ← Compute Marginal
                               Energy( $j$ ) > 0
                               as per the Equation (9)
10:  end for
11:  gini_coefficients ← Compute Gini(marginal_
                               qrs_energy) using the
                               given Equation (10)
12: end for
13: return gini_coefficients

```

the annotated labels of the original datasets. Many of these classifiers are widely used in ECG classification and other types of classification problems in various domains. Brief descriptions of these classifiers are discussed below,

1) K-NEAREST NEIGHBOR (KNN) CLASSIFIER

KNN classifier is a supervised and non-parametric machine learning algorithm that functions without making any prior assumptions about the underlying data, making it particularly well suited for complex, nonlinear classification tasks [61]. Unlike parametric models, KNN classifies data points exclusively based on the proximity of similar instances in the feature space. The ultimate outcomes of the test instances are decided by the majority voting scheme among their closest “k” neighbors, based on distance measures like Euclidean, Manhattan, and Chebyshev measures. In this work, the k value of 3 is selected and the default values are used for the rest of the parameters. The default distance metric is Euclidean distance. Model transparency and simplicity are the main benefits of the KNN classifier. However, a notable drawback is it requires the retention of all training data, demanding substantial memory resources when the dataset is huge. KNN is widely adopted in biomedical applications. Studies such as this have demonstrated the use of KNN to classify various ECG signals. The straightforward and reliable design of KNN on similarity-based classification enables it to accurately distinguish between various cardiac disorders.

2) SUPPORT VECTOR MACHINE (SVM) CLASSIFIER

SVM is a non-probabilistic and linear classification model widely applied in binary classification tasks [62]. It attains classification by creating a hyper-plane that effectively separates data points into distinct classes. Support Vector Machine (SVM) represents each feature point in an N-dimensional space and identifies an optimally positioned hyperplane that would maximize the distance to the closest sample points on both sides from the hyperplane, given a training set of N-dimensional features. This process strengthens the classifier's robustness in handling noisy data, as it aims to hold new, unseen data points correctly classified even if they fall close to the boundary. The data points within the closest premise of the hyperplane are typically known as support vectors, which are crucial for determining the appropriate separation boundary. If the dataset is complex and nonlinear, the SVM can introduce nonlinearity through various kernel functions such as Quadratic, Cubic, or Gaussian, allowing the data to be projected in a multidimensional space, where it becomes linearly separable. SVM has been widely used for many years in various ECG-related applications to detect anomalies. Researchers have been able to accurately detect the presence of anomalies that might otherwise be indistinguishable in a lower dimensional space. We have used the default linear kernel as the kernel function in this work.

3) RANDOM FOREST (RF) CLASSIFIER

Random Forest is one of the ensemble-based ML techniques that builds several decision trees to enhance classification accuracy and minimize the overfitting [63]. A single decision tree is prone to noise and model overfitting when it deals with complex datasets. The forest is constructed by training each tree on a randomly selected subset of features. This bagging or bootstrap aggregation mechanism is used to capture various patterns in the data. The class for a given input is predicted by each decision tree independently. Consequently, the final classification is ascertained with a majority vote among all trees, making the model robust and less sensitive to individual outliers. Furthermore, at each of the trees, only a random subset of selected features is considered for splitting, adding an extra layer of randomness, and this dual randomness will help to minimize the chances for overfitting, even if the model is applied to high-dimensional data. Feature importance metrics could be used for the model's interpretability. Random Forest is widely used in ECG classification and various other biomedical tasks due to its balance of accuracy, robustness, and model interpretability. We have used default values for the applicable hyper-parameters in this work.

4) EXTREME GRADIENT BOOSTING (XGBOOST) CLASSIFIER

XGBoost is a tree-based ensemble method with gradient boosting to attain an enhanced performance in both classification and regression tasks [64]. XGBoost iteratively builds an ensemble of trees where each tree rectifies the errors of its predecessors. This sequential gradient boosting approach

enables XGBoost to address the complexity and the non-linearity in the data. The model's complexity is controlled by the Lasso (L1) and Ridge (L2) regularization parameters which reduce the risks of overfitting and improving generalization. XGBoost uses parallel processing to optimize computational efficiency and uses tree-pruning to handle sparse data. An objective function is used by this classifier. It is a combination of a loss function and a regularization term to lead an accurate and stable model. XGBoost has been successfully utilized in many biomedical applications such as CVD prediction, cancer prediction, gene expression analysis, etc. Model interpretability using feature importance metrics, speed, and accuracy are the major positive factors of XGBoost. Conversely, one of the main drawbacks of this classifier is better performances could be achieved for large and multi-dimensional datasets. Default values are used for the applicable hyper-parameters for our work.

5) GAUSSIAN NAÏVE BAYES (GNB) CLASSIFIER

GNB is a probabilistic classification technique that employs Bayes theorem and presumes that features adhere to a Normal (Gaussian) distribution [65]. Each feature is assumed to contribute independently to the final classification outcome. This algorithm calculates the probability of each class using the Gaussian probability density function to estimate the likelihood of each feature belonging to each class. GNB assigns a test instance to the highest posterior probability class. The algorithm's simplicity allows it to handle larger datasets efficiently and helps with fast predictions. GNB has been applied in ECG classification, genomics, and other bioinformatics tasks. Additionally, GNB provides some level of interpretability by inspecting the feature's likelihood contribution to the prediction. Normally distributed features make GNB a valuable classifier in biomedical data analysis and diagnostics. In this work, default values are used for the relevant hyperparameters analogous to the preceding classifiers.

6) CATEGORICAL BOOSTING (CATBOOST) CLASSIFIER

CatBoost is one of the supervised gradient-boosting algorithms widely applicable in classification and regression studies [66]. It builds an ensemble of several decision trees, optimizing a gradient-boosting mechanism to minimize a loss function. CatBoost is particularly popular in handling structured data. Its effective handling of overfitting renders it suitable for biomedical applications characterized by high-dimensional or imbalanced datasets. CatBoost enhances prediction speed and model interpretability through the utilization of techniques such as oblivious trees (symmetric trees), rendering it especially advantageous for applications such as disease risk prediction, patient classification, and genomic analysis. Another significant advantage of CatBoost is its capability to handle smaller datasets, a prevalent difficulty in biomedical research where data acquisition is limited. Although CatBoost is proficient in handling categorical data,

its complexity may render it less intuitive for interpretation as opposed to more straightforward and simpler models.

7) LINEARBOOST (LB) CLASSIFIER

LinearBoost algorithm, is another variant of boosting methods that combines multiple weak linear models to build a stronger predictive model [67]. Instead of using multiple decision trees as base models, LB uses multiple linear classifiers as underlying base models. This method is particularly effective for high-dimensional data, where linear relationships are dominant, and a series of boosted linear models can capture the underlying patterns without overfitting. LinearBoost is computationally efficient and less resource intensive as opposed to some other complex models. It has been reported that the LinearBoost algorithm outperforms some other prominent Gradient Boosting Decision Tree-based classifiers such as CatBoost, XGBoost, and LightGBM after being tested on a few benchmark datasets. LB has been applied to some biomedical tasks such as genomic studies. To the best of our understanding, this is the first occurrence of LinearBoost being applied in ECG signal classification.

8) LOGISTIC REGRESSION CLASSIFIER

Logistic Regression is also a supervised ML algorithm which is commonly utilized in biomedical and ECG applications [68]. It is commonly applied to binary classification tasks. The probability of the test instance that belongs to a particular class is estimated based on the provided feature vector. Either the sigmoid or logistic function is used to project the predicted values into a probability range from 0 to 1. It seeks to find an optimum fitting model that delineates the links between a collection of independent variables and the binary outcome variable. This property is particularly useful when interpreting the model. Each coefficient delineates the association between a feature and the probability of the target class, rendering it a popular option for healthcare and biomedical research for the interpretability of the results. While logistic regression is simple and interpretable, it presumes a linear relation between input features and the logarithmic transformation of outcomes, which may no longer be applicable in the presence of more intricate and nonlinear interactions. However, Logistic Regression remains a robust and effective choice for classification tasks in biomedical research, where the interpretability of the decision and the model simplicity are priorities.

IV. RESULTS AND DISCUSSIONS

Training and evaluation processes were performed on 88 individual records by combining two different gold-standard public ECG databases as described earlier. We have used 80% of the dataset for training the intelligent classifiers and the remaining 20% of the dataset is for testing the classification models. Further, the “subject-specific” scheme as suggested by the Association for the Advancement of Medical Instrumentation (AAMI) is used to divide the dataset for training and testing. This partition ensures the interpatient separation

which implies that the data corresponds to the same individual is not present in both the training set and the testing set to form an unbiased model. These ECG databases are among a few that contain longer than 30 minutes in duration for each of the individual records. Our experiments were performed for 30 minutes duration of ECG data. Hence, the total number of accounted heartbeats is 183,000 and this is the first time such a larger number of beats are accounted for conventional machine learning-based studies for longer duration ECG signals. TABLE 1 summarizes the accounted number of beats and the analyzed durations of the normal and abnormal ECG signals. Longer duration analysis will help to prevent the subjects from misdirection or misinterpretation by the medical professionals. People will worry unnecessarily if the result is a false positive, and they will not consider it seriously if the result is a false negative. Longer duration analysis is also crucial for detecting a few specific types of abnormal conditions. Moreover, a few minutes from the initial period of the ECG data are unreliable due to several factors such as patient’s fear around the testing equipment, particularly for first-time users, body movements of the patients, and ECG artifacts. Such effects will not have significant impacts while considering the longer-duration signals. Most of the reported studies incorporate very short durations of ECG signals which are even less than a single minute. Additionally, beat-wise classification, either normal or abnormal in a single database may introduce uncertainty in the medical process. For instance, how can a decision be made for a one-minute, or a several second duration ECG recording that indicates half of abnormal and half of normal beats (50% normal) using single beat analysis approaches?

All the experiments were performed on a single lead since the Fantasia-Normal database has only a single lead for ECG signals. The other two leads are, one for respiratory signal and the other is Blood Pressure waveform, meanwhile MIT-BIH Arrhythmia database has two ECG leads. First ECG lead from all the MIT-BIH Arrhythmia records and the first 30 minutes duration of the Fantasia database were selected for the subsequent experiments.

In signal processing applications, energy concentration within time-frequency space plays a crucial role in diagnostic tasks, feature reconstruction and extraction, and classification tasks particularly for non-stationary and multi-component signals. As a result, when using energy concentration of a signal as a feature, it becomes essential to choose a TFD that maximizes the energy concentration for the selected signal. Biomedical signals are vulnerable to several forms of noise during the data-gathering phase. These artifacts can result from poor electrode contact with the body, external interferences such as electrical power noise, or breathing movements. Utilization of TFD for biosignal studies could help mitigate such potential artifacts and enhance the accuracy in detecting abnormal conditions.

According to the existing literature, joint time-frequency analysis methods have performed very well in detecting the abnormalities in ECG signal analysis, seismic signal

TABLE 1. Summary of the databases analyzed in the proposed experiments.

ECG Category	Number of Records	Analyzed Duration	Total Analyzed	Total Number of
		per Record (min)	Duration (min)	Analyzed Beats
Normal or Healthy	40	30	1,200	109,494
Abnormal (Arrhythmia condition)	48	30	1,440	73,510
Total	88	30	2,640	183,004

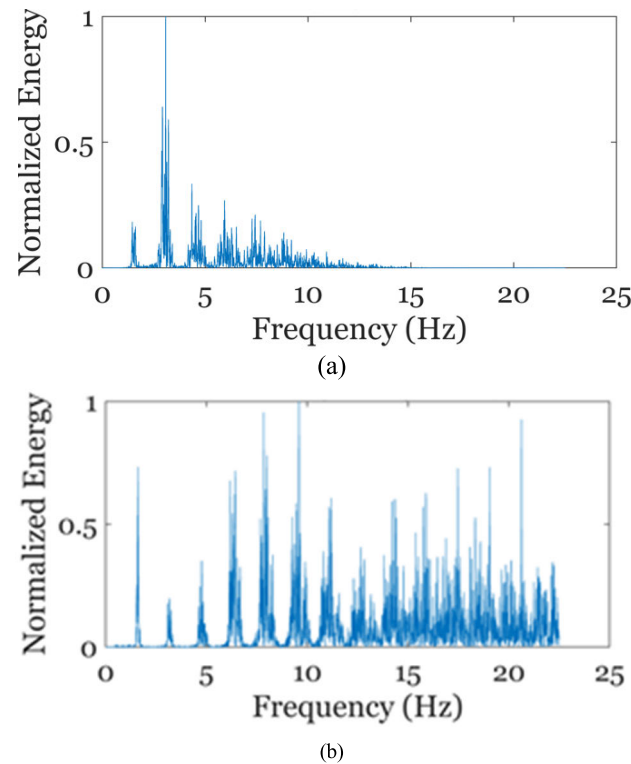


FIGURE 5. Energy concentration variation over the range of frequencies for (a) abnormal ECG for (b) normal ECG.

studies, and machinery fault analysis despite very limited studies being conducted compared to time-domain and frequency-domain approaches. Meanwhile, the CWD method performed well compared to the other Cohen's class of time-frequency methods due to its less susceptible nature to cross-term interference and noisy components that are present in the signals. The Gini index is considered as an exceptional statistical sparsity index to identify the inequalities or dispersions among a particular distribution. This metric has been successfully applied in non-stationary and multi-component signal applications such as machinery fault diagnosis and radar communication. However, this is the first time, the Gini index has been applied on time-frequency energy features in ECG signal analysis. Except the works provided in [69] and [70], Gini index values were derived

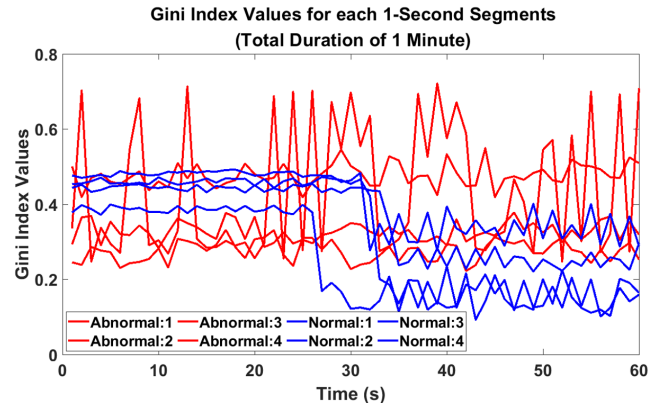


FIGURE 6. Variation of Gini Index measures on 1-second segmentation over the total duration of 60 seconds.

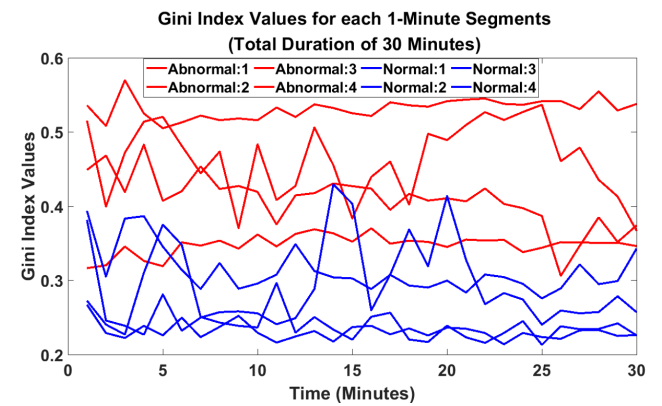


FIGURE 7. Variation of Gini Index measures on 1-minute segmentation over the total duration of 30 minutes.

only from the time domain features in all previously reported signal processing applications.

It is clearly visible from Figure 5 that energy level variations associated with the normal and abnormal ECG signals in each frequency contain meaningful information derived from the signal energy spectrum within the QRS segments. Such information is the reflection of various cardiac activities that can vary at any moment. The QRS wave segments are known to be the most informative region of the ECG signal, as it comprises valid information related to the left-right ventricle depolarizations.

Gini Index measures show commendable performance in classifying the abnormal and normal ECG signals significantly as illustrated in Figure 6 and Figure 7. While the time-window size is increased from 1 second to 1 minute, the distinct difference between abnormal ECG and normal ECG increases in most of the one-minute windows throughout the whole 30-minute duration. This indicates that variations in QRS complex signal energy are higher in abnormal ECG signals as opposed to normal ECG signals.

Detecting Normal and Abnormal ECG is a classification problem in the wide spectrum of machine learning applications. In our study, the performance and effectiveness of each intelligent classifier is assessed using the recommended

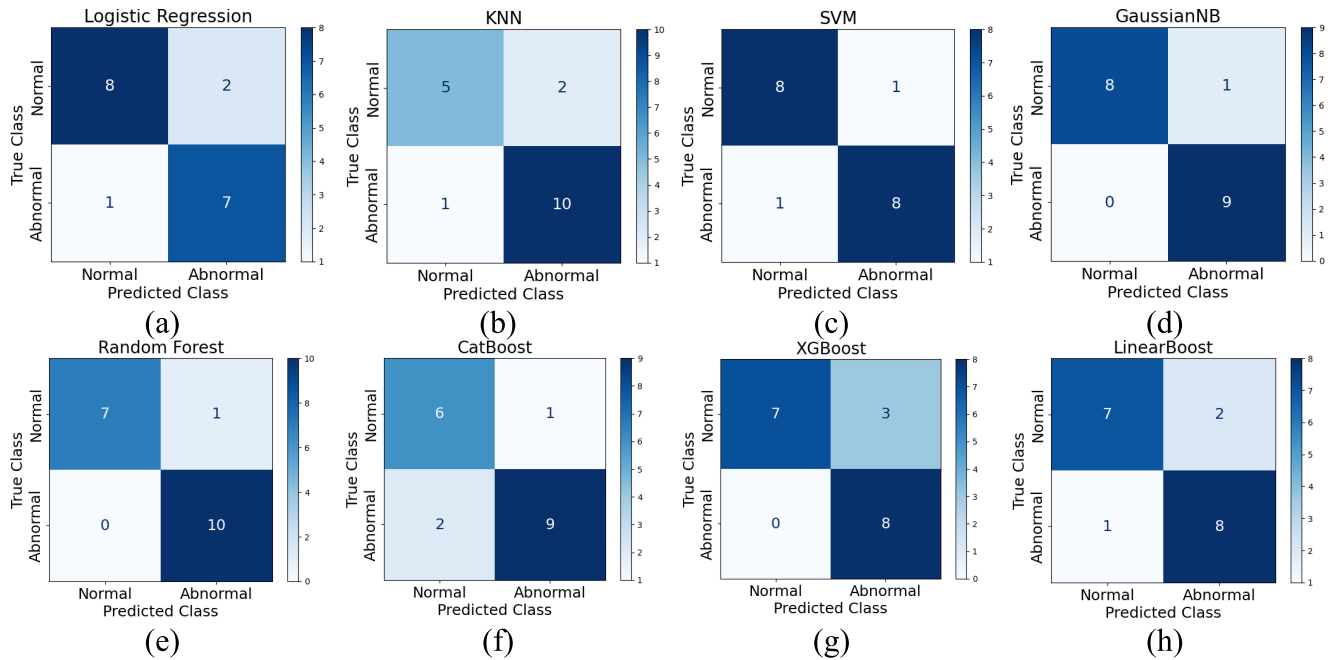


FIGURE 8. Confusion Matrices for all Intelligent Classifiers on Test Data (a). Logistic Regression, (b). K-Nearest Neighbors, (c). Support Vector Machine, (d). Gaussian Naïve Bayes, (e). Random Forest, (f). CatBoost, (g). XGBoost, and (h). LinearBoost respectively.

TABLE 2. Performance metrics of all eight intelligent classifiers on the testing dataset. highlighted in bold entries are the highest obtained values.

Intelligent Classifier	Accuracy (%)	Precision (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
Logistic Regression	83.33	77.78	87.50	80.00	82.35
KNN	83.33	83.33	90.91	71.43	86.96
SVM	88.89	88.89	88.89	88.89	88.89
Gaussian NB	94.44	90.00	100.00	88.89	94.74
Random Forest	94.44	90.00	100.00	87.50	95.24
CatBoost	83.33	90.00	81.82	85.71	85.71
XGBoost	83.33	72.73	100.00	70.00	84.21
LinearBoost	83.33	80.00	88.89	77.78	84.21

metrics from the AAMI. These metrics include Confusion Matrix, Accuracy (Acc), Precision (Pre), Sensitivity (Sen) or Recall, Specificity (Spe), F1-Measure, and Area Under the Curve (AUC). Here, abnormal occurrences are classified as positive, and normal occurrences are classified as negative. Hence, Accuracy denotes the proportion of correctly diagnosed normal and abnormal events out of the total, Precision describes the correctly classified abnormal events that are classified as abnormal, Sensitivity (Recall) reflects the rate of correctly identifying abnormal events, and Specificity shows the rate at which normal events are accurately classified as normal. The harmonic-mean between precision and sensitivity is represented by the F1 measure.

These measures are derived from the confusion matrix as described in the following mathematical expressions (Equations 11-15). AUC denotes the area beneath the Receiver Operating Characteristics (ROC) curve that visualizes the association between the True Positive Rate (Sensitivity) and the True Negative Rate (1-Specificity).

AUC = 1 denotes an ideal classifier, AUC = 0.5 (50%) denotes the model which is not superior to a random guess, and AUC < 0.5 suggests poor performance. Figure 8 demonstrates the confusion matrix for all eight intelligent classifiers for the testing dataset.

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (11)$$

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (13)$$

$$Specificity = \frac{TN}{TN + FP} \quad (14)$$

$$F1 - Measure = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (15)$$

From the medical and healthcare point of view, the sensitivity measure is more critical when compared to accuracy and other metrics as it will help to find out the rate of disease cases that are misclassified as healthy cases which are more dangerous than that of healthy cases are misclassified into disease or abnormal category. This can be identified from the number of TP and FN instances. This also complements the existing practice of an expert-in-the-loop approach where a medical expert provides the diagnosis with the help of technology. Precision and Specificity metrics are influenced by the number of FPs, whereas further medical screening procedures could assist in corroborating the FP cases. However, high sensitivity rates will ensure that most of the abnormal cases are detected at the initial screening process accurately and subsequent treatments can significantly improve the outcomes. This can help to avoid serious health consequences, treatment delays, and life-threatening situations. TABLE 2 provides a

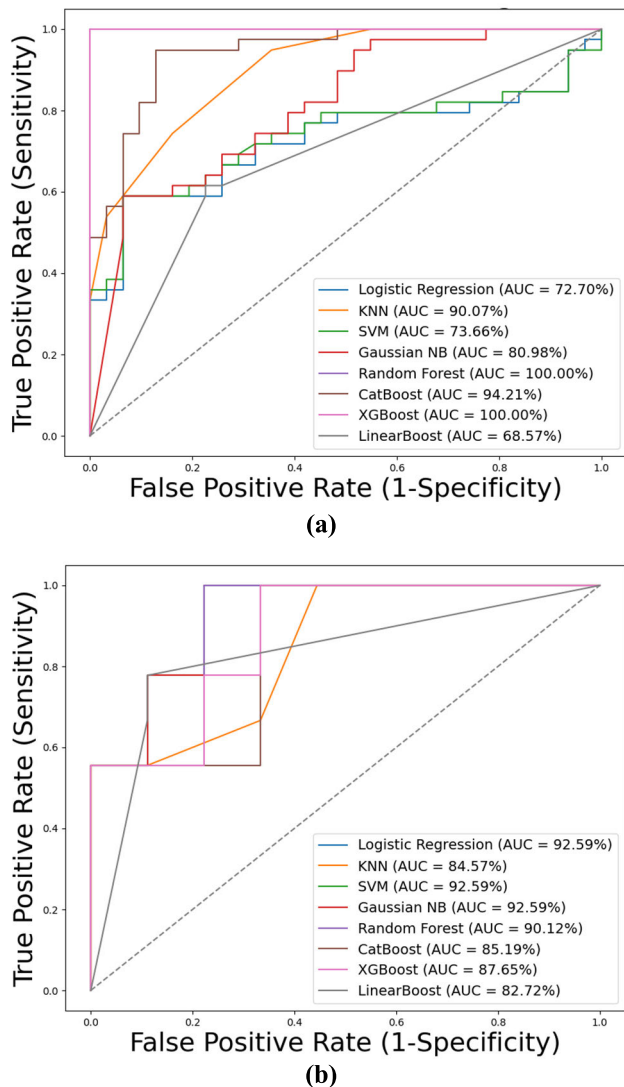


FIGURE 9. Comparative Analysis of ROC Curves for (a). Training Dataset and (b). Test Dataset from Eight Machine Learning Classifiers: Logistic Regression, K-Nearest Neighbor, Support Vector Machine, Gaussian Naïve Bayes, Random Forest, CatBoost, XGBoost, and LinearBoost. The AUC metric is used to assess the effectiveness of each intelligent classifier, with a higher AUC value indicating better performance.

summary of the assessment metrics of each machine learning classifier. It can be noticed from TABLE 2 that three of our intelligent classifiers: Random Forest, Gaussian Naïve Bayes, and XGBoost have achieved a remarkable 100% sensitivity on the evaluated 20% of the testing records. Further to that, Random Forest and Gaussian Naïve Bayes models have outperformed the remaining six classifiers. The Random Forest has achieved 94.44% accuracy, 90% precision, 87.50% specificity, and 95.24% F1-score. Similarly, Gaussian Naïve Bayes has achieved 94.44% accuracy, 90% precision, 88.89% specificity, and 94.74% F1-score respectively.

Moreover, the ROC curves in Figure 9 illustrate the curves for training dataset (80%) and testing dataset (20%) separately. Based on the AUC values of the curves, Random Forest and Gaussian Naïve Bayes precede other classifiers with AUC values of 90.12% and 92.59% on the testing dataset

respectively. Even though Logistic Regression and Support Vector Machines resulted in higher values for AUC of 92.59% on the testing set, they have resulted in low AUC values on training datasets with 72.70% and 73.61% respectively. Further, their accuracy, precision, specificity, and F1-score values are relatively lower compared to those achieved by the Random Forest and the Gaussian Naïve Bayes classifiers.

Our best models (RF and GNB) have achieved 100% sensitivity as they effectively identified all the positive cases due to the distinctive nature of Gini Index features, ensuring no false negatives. The 90% precision was due to some false positives, potentially caused by borderline cases. Accuracy of 94.4% was due to the same false positives. The F1-score of 95.24% was due to the balance between precision and sensitivity. The AUC value of 92.59% demonstrated robust discrimination between healthy and CVD cases across different thresholds.

Random Forest has outperformed several other conventional machine learning classifiers more often in ECG signals and other biomedical classification tasks as reported in the literature. Random Forest is a common choice in many winning real-world data analysis competitions [71]. This could be due to various possible reasons such as it is one of the ensemble approaches consisting of multiple decision trees, which reduces the overfitting and achieves higher accuracy values. Each single tree will be trained on a randomly selected subset of features and data. Further, it can capture complex, nonlinear as well as linear relationships among the features. Hence, it is not required to make a few assumptions on the data like linearity, normality, or homoscedasticity. Random Forest is also robust to noisy and imbalanced data. On top of this award and accolade-winning classifier, the final performance would become explosive when feeding with the powerful and novel TFD-GI features.

Conversely, Gaussian Naïve Bayes will perform exceptionally well for Gaussian or Normally distributed data. Hence, we performed the Shapiro-Wilk normality test on each feature vector separately. The Shapiro-Wilk test evaluates the null hypothesis which assumes the data samples come from a Gaussian-distributed population. A p-value that is less than a selected alpha level ($\alpha = 0.05$) will recommend the rejection of null hypothesis. Meanwhile, A p-value which is greater than $\alpha = 0.05$ will suggest that the data is normally distributed. Among the total of 88 records, 59 feature vectors have resulted in supporting of normality hypothesis with $p > 0.05$, And 29 feature vectors are against the null hypothesis assumption with $p < 0.05$. This is illustrated in Figure 10. Hence, most of the feature vectors are normally distributed, this could be a potential reason for Gaussian Naïve Bayes classifier performed well in the proposed work.

Our model has exceptionally performed well in distinguishing abnormal (arrhythmic) and normal ECG conditions in the heart comparable to other best-performing state-of-the-art techniques reported in existing literature. Notably, our model has outperformed other techniques in sensitivity measure by 100%, which is a crucial factor in medical and

TABLE 3. Comparison of different Methodologies between our work and some of the existing relevant state-of-the-art techniques on Arrhythmia related anomaly detection.

Ref.	Year	Approach	Classification Category	Databases	Number of Records
[21]	2017	Time-domain analysis and DL	ARR, NSR	MIT-BIH ARR	6 Records
[22]	2023	STFT and DL	ARR, NSR, CHF	MIT-BIH ARR, MIT-BIH NSR, and BIDMC	30 ARR, 30 CHF 30 NSR
[26]	2020	STFT, DL, and ML	ARR, NSR, CHF	MIT-BIH ARR, MIT-BIH NSR, and BIDMC	96 ARR, 36 NSR 30 CHF
[72]	2018	STFT, DL, and ML	ARR, NSR	MIT-BIH AF, MIT-BIH VF, EU ST-T, and MIT-BIH NSR	24 AF, 22 VF, 90 ST, 18 NSR
[73]	2020	Fourier-Bessel Transform and DL	ARR, NSR	MIT-BIH ARR	48 Records
[5]	2020	Fusion with wavelet and AR features, and ML	ARR, CHF, NSR	MIT-BIH ARR, MIT-BIH NSR, and BIDMC	30 ARR, 30 CHF 30 NSR
[74]	2016	Wavelet coherence, AR Modelling, and ML	Abnormal, Normal	MIT-BIH NSR, PTB	18 Normal 26 Abnormal
[75]	2015	Time-domain features, PCA, and ML	Abnormal, Normal	PTB	194 Abnormal 52 Normal
[76]	2023	3D-Wavelet transformation, Time-domain features, and ML	ARR, NSR	CPSC 2018	88 ARR, 90 NSR
[77]	2021	DWT, ML, and DL	ARR, NSR	Private dataset	53 ARR, 247 NSR
[78]	2020	Time-domain features and ML	Abnormal, Normal	Private dataset	30 Abnormal 30 Normal
[79]	2022	Wavelet features, ML, and DL	ARR, NSR	MIT-BIH ARR	48 Records
[27]	2022	CWT, and DL	ARR, CHF, NSR	MIT-BIH ARR, MIT-BIH NSR, and BIDMC	30 ARR, 30 CHF 30 NSR
[28]	2022	DL	ARR, NSR	MIT-BIH ARR	48 Records
[20]	2017	DL	ARR, NSR	MIT-BIH ARR	48 Records
[32]	2024	Time-domain features and ML	Abnormal, Normal	Private dataset	21 Abnormal 35 Normal
[36]	2018	Time-domain and Frequency-domain features, and ML	ISC, ARR, NSR	MIT-BIH ARR Fantasia Normal European ST-T	36 ISC, 36 ARR 36 NSR
[35]	2017	Time-domain and Frequency domain features, and ML	ISC, ARR, NSR	MIT-BIH ARR Fantasia Normal European ST-T	36 ISC, 36 ARR 36 NSR
[15]	2019	DL	ARR, NSR	Private dataset, CinC 2017	91,232 records
[16]	2017	DL	ARR, NSR	Private dataset	64,121 records
This work	2024	CWD-TFD, Gini Coefficients of TFD, and ML	ARR, NSR	MIT-BIH ARR Fantasia Normal	48 ARR, 40 NSR

DWT: Discrete Wavelet Transform, PCA: Principal Component Analysis, DL: Deep Learning, ML: Machine Learning, AR: Auto Regression, CWT: Continuous Wavelet Transform, STFT: Short Time Fourier Transform, NSR: Normal Sinus Rhythm, CHF: Congestive Heart Failure, ISC: Ischemia, and ARR: Arrhythmia

healthcare settings, where sensitivity is paramount. Further, we have incorporated and analyzed a larger number of heartbeats in the time-frequency domain with a 30-minute duration of two publicly available gold-standard ECG databases. Only a limited number of research studies surpassed the number of beats analyzed in total and their approach was in time-domain and deep learning-based classification, where the interpretability of the diagnosis is limited.

Moreover, they performed a cardiologist level of identification in which they categorize the different types of arrhythmia-related beat types, whereas our study focuses on the initial medical screening process to detect the abnormal conditions present in the heart. Our approach can assist General Practitioners (GPs) who are not certified cardiologists during the initial medical screening process for CVD-related diseases prone to human errors due to the lack of expertise level of GPs compared to the Medical Specialists in that

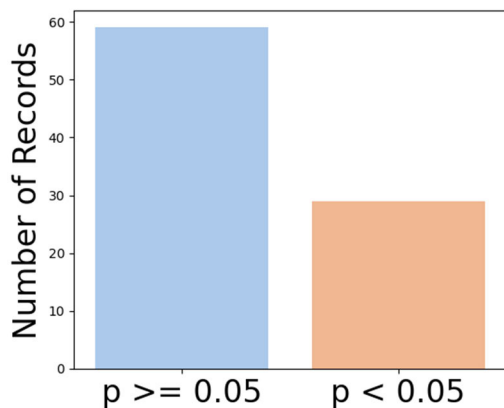
specific field [37]. Table 3 provides a synopsis of different types of approaches used in the ECG signal anomaly detection literature and Table 4 provides a performance comparative analysis summary of our proposed approach with some of the existing state-of-the-art techniques.

However, it is worth highlighting that there is a possibility that some researchers may have selectively chosen data in favor of their algorithms, as certain studies did not utilize all individual records or the whole duration of the signals from the used databases. In work [5], [21], [22], [27], [35], and [36], researchers have not covered all the individual records available in that specific database. Most of the highlighted studies analyzed a portion of the time window of the signal except the works presented in [20], [75], and [77]. Meanwhile, researchers in [35] and [36] did not divide the data into training and testing cohorts, they have just used the complete data exclusively for training

TABLE 4. Comparative analysis of performances between our work and some of the existing relevant state-of-the-art techniques on Arrhythmia related anomaly detection.

Ref.	Total Number of Associated Heartbeats	Performance on Training Dataset	Performance on Testing or Validation Dataset				
		Acc (%)	Acc (%)	Sen (%)	Pre (%)	Spe (%)	F1-Score
[21]	416	98.51	92.00	NR	NR	NR	NR
[22]	Less than 10,000	NR	99.20	99.20	99.20	99.60	99.20
[26]	Less than 35,000	NR	96.77	98.33	94.33	NR	96.00
[72]	7008	NR	97.23	97.02	97.76	NR	97.35
[73]	2880	NR	90.07	92.12	NR	88.10	90.07
[5]	Approx. 48,000	NR	93.33	NR	NR	NR	NR
[74]	Approx. 7,000	NR	99.1	96.97	NR	99.43	NR
[75]	Approx. 30,000	NR	88.52	89.68	NR	84.62	NR
[76]	6400	NR	99.02	98.87	98.90	NR	98.87
[77]	Approx. 60,000	NR	93.00	95.00	93.00	NR	95.00
[78]	Less than 10,000	NR	90.00	90.00	NR	90.00	NR
[79]	84,615	NR	91.92	90.21	NR	95.18	NR
[27]	Less than 5,000	NR	97.30	97.10	98.79	98.50	97.40
[28]	102,179	NR	98.36	94.36	89.40	NR	91.67
[20]	109,449	NR	94.03	96.71	97.86	91.54	97.28
[32]	Approx. 5,000	NR	97.00	95.00	NR	95.00	NR
[36]	10,100	97.85	NR	NR	NR	NR	NR
[35]	10,100	97.03	NR	NR	NR	NR	NR
[15]	Approx. 3.4 million	NR	NR	83.00	84.40	93.27	83.70
[16]	Approx. 2.4 million	NR	NR	82.70	80.90	NR	80.90
This work	183,000	98.57	94.40	100.00	90.00	87.50	95.24

NR: Not Reported.

**FIGURE 10.** Results from Shapiro-Wilk test by p-value category for each feature vector for the total number of records. $P \geq 0.05$ supports the null hypothesis which assumes data come from a normal distribution, whereas $p < 0.05$ will reject the assumption of null hypothesis.

only and reported the accuracy metrics regarding the training dataset.

Unlike traditional closed box models, XAI techniques help health experts understand why a model predicts certain types of abnormalities, improving both diagnostic confidence and patient outcomes. The concept of explainability is not new; it has evolved from conventional probability-based models, rule-based models, decision trees to the recent explainable AI tools such as LIME, SHAP, and Grad-CAM. Four levels of explainability are applied for AI-ECG applications: 1. For justifying the decisions, 2. For enhancing the model performance, 3. For controlling the decisions, 4. For discovering and learning new patterns [18]. Our proposed approach falls under model explainability for justifying the obtained

decisions. Most of the previously reported XAI works can be summarized into the following explainability techniques: intrinsic or post-hoc analysis, Global or Local, model-specific or model-agnostic. A detailed comparison of applied explainability techniques of previous works can be found in Table 2 of work [19]. Except KNN and SVM, most of the classifiers employed in this study are decision tree-based models, while the Naïve Bayes classifier is based on probabilistic principles. Notably, all these classifiers fall under the category of XAI techniques. Moreover, the features represented by the Gini Index measures of non-linear time-frequency energy concentration variations depicted in Figure 6 and Figure 7 would also be helpful in model explainability for justifying the diagnostic decisions. However, unlike the previously reported works in Table 2 of [19], our work has a combination of Intrinsic-Global-Model Agnostic nature of model explainability. Intrinsic nature is due to the simple internal structure and decision-making process without the need for additional explanation techniques. The globality nature is achieved by the underlying Gini Index features that have a significant impact on its predictions across the entire dataset rather than focusing on individual data points. Model agnostic is independent of specific architecture or internal workings of the model. Hence, this work is a unique contribution to model explainability compared to the earlier reported works.

V. CONCLUSION AND FUTURE DIRECTION

Conventionally, ECG signal classification is performed in two ways. One approach is classifying the heartbeats exclusively into specific types of Arrhythmias such as Ventricular tachycardia, Atrial flutter, Bradycardia, Atrial fibrillation,

Supraventricular tachycardia, Ventricular fibrillation, Premature heartbeat, Supraventricular arrhythmias, Premature contraction, and Normal heartbeats. This is referred to as cardiologist-level detection which reflects a medical expert. The second approach is classifying the patients or subjects whether they have any types of cardiovascular disorders, or whether their heart is healthy. This latter approach is very crucial in the initial medical screening stages, which are generally done by a General Physician prior to the case being transmitted to the experts or further advanced screening stages. This process is prone to human errors due to the level of experience and skills of GPs when compared to the level of an expert in that field, who are limited in number. Moreover, GPs focus on a broad spectrum of diseases and are not limited to a single type of illness. Hence, our study proposes an automated solution to this kind of initial screening process due to the significance of CVDs that could support mainly the GPs or could be deployed in remote areas across the world where a very limited number of medical professionals are available or expensive to reach them. On the other hand, smart wearable devices such as smartwatches are limited to detecting a single type of Arrhythmia condition only. Our proposed algorithm could be used with such types of smart devices as well that have now become more pervasive wearable. Moreover, many of the current algorithms including such smart devices are based exclusively on time domain analysis only, which is not capable of capturing the complex, non-stationarity, multi-component, and non-linearity nature of ECG kinds of signals.

Our proposed approach has the following advantages and contributions,

1. Providing a nonlinear time-frequency approach for analyzing the ECG signals that are non-stationary and multicomponent in nature.
2. Longer duration analysis for each subject is about 30 minutes duration and covers a larger number of heartbeats which is important to reduce the detection errors.
3. Proposing a novel method of extracting Gini index features from the Nonlinear TFD energy contents of ECG signals.
4. These Gini index measurements of energy content could be helpful in explaining or interpreting the final diagnostic results.
5. Providing a comparative assessment of various intelligent classifiers and their effect on the final classification performance.
6. Obtained a 100% sensitivity rate which is more advantageous in medical settings. Other metrics are also commendable in comparison to the current state-of-the-art approaches (Acc = 94.40%, Pre = 90.00, F1-Score = 95.24%, and AUC = 92.59%)
7. Our algorithm can be integrated into any kind of wearable device, web-based platform, and any other type of offline-online monitoring system, due to less data-intensive machine learning approach compared to deep learning.

One of the challenges in this approach is the requirement of high memory and processing for nonlinear time-frequency analysis. However, this is negligible in front of the exponential growth in computing resources in recent times including cloud-distributed based high-performance computing platforms. This processing part could be executed in the centralized back-end system. Hence, it will not be constrained by the front-end devices. Our research is currently limited to arrhythmia-based abnormality on MIT-BIH Arrhythmia and Fantasia Normal databases. However, longer duration (30 minutes or more) arrhythmia databases are very limited in the public domain. Hence, we are currently in the process of collecting our own arrhythmia-abnormal cases in Malaysia to further validate the model and to ensure the practical applicability, model generalization, and robustness of the proposed approach. Moreover, we expect to incorporate other severe types of cardiac diseases such as Ischemia in the future work. Furthermore, different types of features with similar strengths of the Gini index will be incorporated to enhance the robustness and accuracy of this proposed algorithm.

ACKNOWLEDGMENT

The authors express their gratitude to the Universiti Putra Malaysia and the Embassy of France to Malaysia for funding this research under GP-IPS and MyTiger grant schemes. Authors also acknowledge the IEEE Access reviewers for their constructive comments and feedback on enhancing the quality of this manuscript.

REFERENCES

- [1] M. Gu, Y. Zhang, Y. Wen, G. Ai, H. Zhang, P. Wang, and G. Wang, "A lightweight convolutional neural network hardware implementation for wearable heart rate anomaly detection," *Comput. Biol. Med.*, vol. 155, Mar. 2023, Art. no. 106623, doi: [10.1016/j.combiomed.2023.106623](https://doi.org/10.1016/j.combiomed.2023.106623).
- [2] M. Di Cesare, P. Perel, S. Taylor, C. Kabudula, H. Bixby, T. A. Gaziano, D. V. McGhie, J. Mwangi, B. Pervan, J. Narula, D. Pineiro, and F. J. Pinto, "The heart of the world," *Global Heart*, vol. 19, no. 1, pp. 1–12, Jan. 2024, doi: [10.5334/gh.1288](https://doi.org/10.5334/gh.1288).
- [3] S. M. Dunlay, Q. R. Pack, R. J. Thomas, J. M. Killian, and V. L. Roger, "Participation in cardiac rehabilitation, readmissions, and death after acute myocardial infarction," *Amer. J. Med.*, vol. 127, no. 6, pp. 538–546, Jun. 2014, doi: [10.1016/j.amjmed.2014.02.008](https://doi.org/10.1016/j.amjmed.2014.02.008).
- [4] G. Y. H. Lip, D. A. Lane, R. Lenarczyk, G. Boriani, W. Doehner, L. A. Benjamin, M. Fisher, D. Lowe, R. L. Sacco, R. Schnabel, C. Watkins, G. Ntaios, and T. Potpara, "Integrated care for optimizing the management of stroke and associated heart disease: A position paper of the European society of cardiology council on stroke," *Eur. Heart J.*, vol. 43, no. 26, pp. 2442–2460, Jul. 2022, doi: [10.1093/eurheartj/ehac245](https://doi.org/10.1093/eurheartj/ehac245).
- [5] S. Nahak and G. Saha, "A fusion based classification of normal, arrhythmia and congestive heart failure in ECG," in *Proc. Nat. Conf. Commun. (NCC)*, Feb. 2020, pp. 1–6, doi: [10.1109/NCC48643.2020.9056095](https://doi.org/10.1109/NCC48643.2020.9056095).
- [6] B. F. Cuneo et al., "Home monitoring for fetal heart rhythm during anti-ro pregnancies," *J. Amer. College Cardiol.*, vol. 72, no. 16, pp. 1940–1951, Oct. 2018, doi: [10.1016/j.jacc.2018.07.076](https://doi.org/10.1016/j.jacc.2018.07.076).
- [7] D. Jakubovitz, R. Giryas, and M. R. D. Rodrigues, "Generalization error in deep learning," in *Proc. 3rd Int. MATHEON Conf.*, H. Boche, G. Caire, R. Calderbank, G. Kutyniok, R. Mathar, and P. Petersen, Eds., Cham, Switzerland: Springer, 2019, pp. 153–193, doi: [10.1007/978-3-319-73074-5_5](https://doi.org/10.1007/978-3-319-73074-5_5).
- [8] A. F. Hussein, S. J. Hashim, F. Z. Rokhani, and W. A. Wan Adnan, "An automated high-accuracy detection scheme for myocardial ischemia based on multi-lead long-interval ECG and Choi-Williams time-frequency analysis incorporating a multi-class SVM classifier," *Sensors*, vol. 21, no. 7, p. 2311, Mar. 2021, doi: [10.3390/s21072311](https://doi.org/10.3390/s21072311).

- [9] A. F. Hussein, S. J. Hashim, A. F. A. Aziz, F. Z. Rokhani, and W. A. W. Adnan, "Performance evaluation of time-frequency distributions for ECG signal analysis," *J. Med. Syst.*, vol. 42, no. 1, p. 15, Jan. 2018, doi: [10.1007/s10916-017-0871-8](#).
- [10] R. J. Martis, U. R. Acharya, and H. Adeli, "Current methods in electrocardiogram characterization," *Comput. Biol. Med.*, vol. 48, pp. 133–149, May 2014, doi: [10.1016/j.compbiomed.2014.02.012](#).
- [11] Y. Miao, M. Zhao, and J. Hua, "Research on sparsity indexes for fault diagnosis of rotating machinery," *Measurement*, vol. 158, Jul. 2020, Art. no. 107733, doi: [10.1016/j.measurement.2020.107733](#).
- [12] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 6, pp. 862–871, Jun. 1989.
- [13] F. Hlawatsch and F. Auger, *Time-Frequency Analysis*. Hoboken, NJ, USA: Wiley, 2013.
- [14] R. Herman, A. Demolder, B. Vavrik, M. Martonak, V. Boza, V. Kresnakova, A. Iring, T. Palus, J. Bahyl, O. Nelis, M. Beles, D. Fabbriatore, L. Perl, J. Bartunek, and R. Hatala, "Validation of an automated artificial intelligence system for 12-lead ECG interpretation," *J. Electrocardiol.*, vol. 82, pp. 147–154, Jan. 2024, doi: [10.1016/j.jelectrocard.2023.12.009](#).
- [15] A. Y. Hannun, P. Rajpurkar, M. Haghighpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Med.*, vol. 25, no. 1, pp. 65–69, Jan. 2019, doi: [10.1038/s41591-018-0268-3](#).
- [16] P. Rajpurkar, A. Y. Hannun, M. Haghighpanahi, C. Bourn, and A. Y. Ng, "Cardiologist-level arrhythmia detection with convolutional neural networks," 2017, *arXiv:1707.01836*.
- [17] L. Nanni, S. Ghidoni, and S. Brahnam, "Handcrafted vs. non-handcrafted features for computer vision classification," *Pattern Recognit.*, vol. 71, pp. 158–172, Nov. 2017, doi: [10.1016/j.patcog.2017.05.025](#).
- [18] Z. Sadeghi, R. Alizadehsani, M. A. Cifci, S. Kausar, R. Rehman, P. Mahanta, P. K. Bora, A. Almasri, R. S. Alkhalwaldeh, S. Hussain, B. Alatas, A. Shoeibi, H. Moosaei, M. Hladik, S. Nahavandi, and P. M. Pardalos, "A review of explainable artificial intelligence in healthcare," *Comput. Electr. Eng.*, vol. 118, Aug. 2024, Art. no. 109370, doi: [10.1016/j.compeleceng.2024.109370](#).
- [19] S. Bharati, M. R. H. Mondal, and P. Podder, "A review on explainable artificial intelligence for healthcare: Why, how, and when?" *IEEE Trans. Artif. Intell.*, vol. 5, no. 4, pp. 1429–1442, Apr. 2024, doi: [10.1109/TAI.2023.3266418](#).
- [20] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, M. Adam, A. Gertych, and R. S. Tan, "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389–396, Oct. 2017, doi: [10.1016/j.compbiomed.2017.08.022](#).
- [21] A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning," *Proc. Comput. Sci.*, vol. 120, pp. 268–275, Jan. 2017, doi: [10.1016/j.procs.2017.11.238](#).
- [22] Y. D. Daydulo, B. L. Thamineni, and A. A. Dawud, "Cardiac arrhythmia detection using deep learning approach and time frequency representation of ECG signals," *BMC Med. Informat. Decis. Making*, vol. 23, no. 1, p. 232, Oct. 2023, doi: [10.1186/s12911-023-02326-w](#).
- [23] B. Hou, J. Yang, P. Wang, and R. Yan, "LSTM-based auto-encoder model for ECG arrhythmias classification," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 4, pp. 1232–1240, Apr. 2020, doi: [10.1109/TIM.2019.2910342](#).
- [24] Ö. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Comput. Biol. Med.*, vol. 96, pp. 189–202, May 2018, doi: [10.1016/j.compbiomed.2018.03.016](#).
- [25] H. M. Lynn, S. B. Pan, and P. Kim, "A deep bidirectional GRU network model for biometric electrocardiogram classification based on recurrent neural networks," *IEEE Access*, vol. 7, pp. 145395–145405, 2019, doi: [10.1109/ACCESS.2019.2939947](#).
- [26] A. Çınar and S. A. Tuncer, "Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks," *Comput. Methods Biomechanics Biomed. Eng.*, vol. 24, no. 2, pp. 203–214, Jan. 2021, doi: [10.1080/10255842.2020.1821192](#).
- [27] P. Madan, V. Singh, D. P. Singh, M. Diwakar, B. Pant, and A. Kishor, "A hybrid deep learning approach for ECG-based arrhythmia classification," *Bioengineering*, vol. 9, no. 4, p. 152, Apr. 2022, doi: [10.3390/bioengineering9040152](#).
- [28] S. Bhatia, S. K. Pandey, A. Kumar, and A. Alshuhail, "Classification of electrocardiogram signals based on hybrid deep learning models," *Sustainability*, vol. 14, no. 24, p. 16572, Dec. 2022, doi: [10.3390/su142416572](#).
- [29] O. D'angelis, L. Bacco, L. Vollero, and M. Merone, "Advancing ECG biometrics through vision transformers: A idence-driven approach," *IEEE Access*, vol. 11, pp. 140710–140721, 2023, doi: [10.1109/ACCESS.2023.3338191](#).
- [30] Y. Tao, B. Xu, and Y. Zhang, "Refined self-attention transformer model for ECG-based arrhythmia detection," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–14, 2024, doi: [10.1109/TIM.2024.3400302](#).
- [31] R. M. Al-Tam, A. M. Al-Hejri, E. Naji, F. A. Hashim, S. S. Alshamrani, A. Alshehri, and S. M. Narangale, "A hybrid framework of transformer encoder and residential conventional for cardiovascular disease recognition using heart sounds," *IEEE Access*, vol. 12, pp. 123099–123113, 2024, doi: [10.1109/ACCESS.2024.3451660](#).
- [32] Z. Alimbayeva, C. Alimbayev, K. Ozhikenov, N. Bayanbay, and A. Ozhikenova, "Wearable ECG device and machine learning for heart monitoring," *Sensors*, vol. 24, no. 13, p. 4201, Jun. 2024, doi: [10.3390/s24134201](#).
- [33] S. M. Malakouti, "Heart disease classification based on ECG using machine learning models," *Biomed. Signal Process. Control*, vol. 84, Jul. 2023, Art. no. 104796, doi: [10.1016/j.bspc.2023.104796](#).
- [34] C. H. Lee and S. H. Kim, "ECG measurement system for vehicle implementation and heart disease classification using machine learning," *IEEE Access*, vol. 11, pp. 17968–17982, 2023, doi: [10.1109/ACCESS.2023.3245565](#).
- [35] A. K. Bhoi, K. S. Sherpa, and B. Khandelwal, "Arrhythmia and ischemia classification and clustering using QRS-ST-T (QT) analysis of electrocardiogram," *Cluster Comput.*, vol. 21, no. 1, pp. 1033–1044, Mar. 2018, doi: [10.1007/s10586-017-0957-6](#).
- [36] A. K. Bhoi, K. S. Sherpa, and B. Khandelwal, "Ischemia and arrhythmia classification using time-frequency domain features of QRS complex," *Proc. Comput. Sci.*, vol. 132, pp. 606–613, Jan. 2018, doi: [10.1016/j.procs.2018.05.014](#).
- [37] R. R. Sharma, M. Kumar, and R. B. Pachori, "Joint time-frequency domain-based CAD disease sensing system using ECG signals," *IEEE Sensors J.*, vol. 19, no. 10, pp. 3912–3920, May 2019, doi: [10.1109/JSEN.2019.2894706](#).
- [38] Z. Abduh, E. A. Nehary, M. A. Wahed, and Y. M. Kadah, "Classification of heart sounds using fractional Fourier transform based mel-frequency spectral coefficients and traditional classifiers," *Biomed. Signal Process. Control*, vol. 57, Mar. 2020, Art. no. 101788, doi: [10.1016/j.bspc.2019.101788](#).
- [39] S. Aziz, S. Ahmed, and M.-S. Alouini, "ECG-based machine-learning algorithms for heartbeat classification," *Sci. Rep.*, vol. 11, no. 1, p. 18738, Sep. 2021, doi: [10.1038/s41598-021-97118-5](#).
- [40] A. J. D. Krupa, S. Dhanalakshmi, and R. Kumar, "Joint time-frequency analysis and non-linear estimation for fetal ECG extraction," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103569, doi: [10.1016/j.bspc.2022.103569](#).
- [41] G. R. Pereira, L. F. de Oliveira, and J. Nadal, "Reducing cross terms effects in the Choi-Williams transform of mioelectric signals," *Comput. Methods Programs Biomed.*, vol. 111, no. 3, pp. 685–692, Sep. 2013, doi: [10.1016/j.cmpb.2013.06.004](#).
- [42] B. Boashash, *Time-Frequency Signal Analysis and Processing: A Comprehensive Reference*. New York, NY, USA: Academic, 2015.
- [43] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, pp. e215–e220, Jun. 2000, doi: [10.1161/01.cir.101.23.e215](#).
- [44] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. Med. Biol. Mag.*, vol. 20, no. 3, pp. 45–50, May 2001, doi: [10.1109/51.932724](#).
- [45] N. Iyengar, C. K. Peng, R. Morin, A. L. Goldberger, and L. A. Lipsitz, "Age-related alterations in the fractal scaling of cardiac interbeat interval dynamics," *Amer. J. Physiol.-Regulatory, Integrative Comparative Physiol.*, vol. 271, no. 4, pp. R1078–R1084, Oct. 1996, doi: [10.1152/ajpregu.1996.271.4.r1078](#).
- [46] E. Ajaraga and M. Gusev, "Analysis of sampling frequency and resolution in ECG signals," in *Proc. 25th Telecommun. Forum (TELFOR)*, Nov. 2017, pp. 1–4, doi: [10.1109/TELFOR.2017.8249438](#).
- [47] M. Malik, "Heart rate variability: Standards of measurement, physiological interpretation, and clinical use: Task force of the European society of cardiology and the North American society for pacing and electrophysiology," *Ann. Noninvasive Electrocardiol.*, vol. 1, no. 2, pp. 151–181, Apr. 1996.

- [48] G. Lenis, N. Pilia, A. Loewe, W. H. W. Schulze, and O. Dössel, "Comparison of baseline wander removal techniques considering the preservation of ST changes in the ischemic ECG: A simulation study," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–13, Mar. 2017, doi: [10.1155/2017/925029](https://doi.org/10.1155/2017/925029).
- [49] J. O. Toole and B. Boashash, "Fast and memory-efficient algorithms for computing quadratic time–frequency distributions," *Appl. Comput. Harmon. Anal.*, vol. 35, no. 2, pp. 350–358, Sep. 2013.
- [50] B. Fatimah, P. Singh, A. Singhal, D. Pramanick, S. Pranav, and R. B. Pachori, "Efficient detection of myocardial infarction from single lead ECG signal," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102678, doi: [10.1016/j.bspc.2021.102678](https://doi.org/10.1016/j.bspc.2021.102678).
- [51] A. Asgharzadeh-Bonab, M. C. Amirani, and A. Mehri, "Spectral entropy and deep convolutional neural network for ECG beat classification," *Bio-cybernetics Biomed. Eng.*, vol. 40, no. 2, pp. 691–700, Apr. 2020, doi: [10.1016/j.bbe.2020.02.004](https://doi.org/10.1016/j.bbe.2020.02.004).
- [52] T. Li and M. Zhou, "ECG classification using wavelet packet entropy and random forests," *Entropy*, vol. 18, no. 8, p. 285, Aug. 2016, doi: [10.3390/e18080285](https://doi.org/10.3390/e18080285).
- [53] S. Ansari, N. Farzaneh, M. Duda, K. Horan, H. B. Andersson, Z. D. Goldberger, B. K. Nallamothu, and K. Najarian, "A review of automated methods for detection of myocardial ischemia and infarction using electrocardiogram and electronic health records," *IEEE Rev. Biomed. Eng.*, vol. 10, pp. 264–298, 2017, doi: [10.1109/RBME.2017.2757953](https://doi.org/10.1109/RBME.2017.2757953).
- [54] C. Sridhar, O. S. Lih, V. Jahmunah, J. E. W. Koh, E. J. Ciaccio, T. R. San, N. Arunkumar, S. Kadry, and U. R. Acharya, "Accurate detection of myocardial infarction using non linear features with ECG signals," *J. Ambient Intell. Humanized Comput.*, vol. 12, no. 3, pp. 3227–3244, Mar. 2021, doi: [10.1007/s12652-020-02536-4](https://doi.org/10.1007/s12652-020-02536-4).
- [55] Y. Miao, J. Wang, B. Zhang, and H. Li, "Practical framework of Gini index in the application of machinery fault feature extraction," *Mech. Syst. Signal Process.*, vol. 165, Feb. 2022, Art. no. 108333, doi: [10.1016/j.ymssp.2021.108333](https://doi.org/10.1016/j.ymssp.2021.108333).
- [56] M. W. Muelas, F. Mughal, S. O'Hagan, P. J. Day, and D. B. Kell, "The role and robustness of the Gini coefficient as an unbiased tool for the selection of Gini genes for normalising expression profiling data," *Sci. Rep.*, vol. 9, no. 1, p. 17960, Nov. 2019, doi: [10.1038/s41598-019-54288-7](https://doi.org/10.1038/s41598-019-54288-7).
- [57] S. Feranchuk, N. Belkova, U. Potapova, D. Kuzmin, and S. Belikov, "Evaluating the use of diversity indices to distinguish between microbial communities with different traits," *Res. Microbiol.*, vol. 169, nos. 4–5, pp. 254–261, May 2018, doi: [10.1016/j.resmic.2018.03.004](https://doi.org/10.1016/j.resmic.2018.03.004).
- [58] G. Dai, R. Li, and S. Ma, "Research on the equity of health resource allocation in TCM hospitals in China based on the Gini coefficient and agglomeration degree: 2009–2018," *Int. J. Equity Health*, vol. 21, no. 1, p. 145, Oct. 2022, doi: [10.1186/s12939-022-01749-7](https://doi.org/10.1186/s12939-022-01749-7).
- [59] B. K. Das and H. S. Dutta, "GFNB: Gini index–based fuzzy naive Bayes and blast cell segmentation for leukemia detection using multi-cell blood smear images," *Med. Biol. Eng. Comput.*, vol. 58, no. 11, pp. 2789–2803, Nov. 2020, doi: [10.1007/s11517-020-02249-y](https://doi.org/10.1007/s11517-020-02249-y).
- [60] M. F. de Godoy, B. A. Rudnick, J. V. de M. Reichert, M. F. de Godoy, B. A. Rudnick, and J. V. de M. Reichert, "Use of the Gini coefficient for the analysis of heart rate variability in sick and healthy individuals," in *Time Series Analysis—Recent Advances, New Perspectives and Applications*. London, U.K.: IntechOpen, 2023, doi: [10.5772/intechopen.1002956](https://doi.org/10.5772/intechopen.1002956).
- [61] A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, "K-nearest neighbor classification," in *Data Mining in Agriculture*, A. Mucherino, P. J. Papajorgji, and P. M. Pardalos, Eds., New York, NY, USA: Springer, 2009, pp. 83–106, doi: [10.1007/978-0-387-88615-2_4](https://doi.org/10.1007/978-0-387-88615-2_4).
- [62] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, pp. 230–267, Jan. 2010, doi: [10.1039/b918972f](https://doi.org/10.1039/b918972f).
- [63] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, Oct. 2001, doi: [10.1023/a:1010933404324](https://doi.org/10.1023/a:1010933404324).
- [64] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," 2016, *arXiv:1603.02754*.
- [65] V. B. Vikramkumar and Trilochan, "Bayes and naive Bayes classifier," 2014, *arXiv:1404.0933*.
- [66] L. Prokhorenkova, G. Gusev, A. Vorobev, A. V. Dorogush, and A. Gulin, "CatBoost: Unbiased boosting with categorical features," in *Proc. Adv. Neural Inf. Process. Syst.*, 2018, pp. 1–14. Accessed: Nov. 25, 2024. [Online]. Available: <https://proceedings.neurips.cc/paper/2018/hash/14491b756b3a51daac41c24863285549-Abstract.html>
- [67] *Welcome To LinearBoost's Documentation!—LinearBoost 0.1 Documentation*. Accessed: Nov. 25, 2024. [Online]. Available: <https://linearboost.readthedocs.io/en/latest/>
- [68] M. P. LaValley, "Logistic regression," *Circulation*, vol. 117, no. 18, pp. 2395–2399, 2008.
- [69] I. Orović, S. Stanković, and M. Beko, "On the use of Gini coefficient for measuring time-frequency distribution concentration and parameters selection," *Math. Problems Eng.*, vol. 2022, Oct. 2022, Art. no. 7731309, doi: [10.1155/2022/7731309](https://doi.org/10.1155/2022/7731309).
- [70] L. Zhili, W. Yunxin, H. Lei, and Z. Jue, "Joint time-frequency analysis of capillary tip vibration in thermosonic wire bonding," in *Proc. Conf. High Density Microsyst. Design Packag. Compon. Failure Anal.*, Jun. 2005, pp. 1–6, doi: [10.1109/hdp.2005.251447](https://doi.org/10.1109/hdp.2005.251447).
- [71] S. J. Rigatti, "Random forest," *J. Insurance Med.*, vol. 47, no. 1, pp. 31–39, Jan. 2017, doi: [10.17849/inm-47-01-31-39.1](https://doi.org/10.17849/inm-47-01-31-39.1).
- [72] M. Salem, S. Taheri, and J. Yuan, "ECG arrhythmia classification using transfer learning from 2- dimensional deep CNN features," in *Proc. IEEE Biomed. Circuits Syst. Conf. (BioCAS)*, Oct. 2018, pp. 1–4, doi: [10.1109/BIOCAS.2018.8584808](https://doi.org/10.1109/BIOCAS.2018.8584808).
- [73] A. Sharma, N. Garg, S. Patidar, R. S. Tan, and U. R. Acharya, "Automated pre-screening of arrhythmia using hybrid combination of Fourier–Bessel expansion and LSTM," *Comput. Biol. Med.*, vol. 120, May 2020, Art. no. 103753, doi: [10.1016/j.combiomed.2020.103753](https://doi.org/10.1016/j.combiomed.2020.103753).
- [74] P. Kora and K. S. R. Krishna, "ECG based heart arrhythmia detection using wavelet coherence and bat algorithm," *Sens. Imag.*, vol. 17, no. 1, p. 12, Jun. 2016, doi: [10.1007/s11220-016-0136-5](https://doi.org/10.1007/s11220-016-0136-5).
- [75] R. Huang and Y. Zhou, "Disease classification and biomarker discovery using ECG data," *BioMed Res. Int.*, vol. 2015, Nov. 2015, Art. no. 680381, doi: [10.1155/2015/680381](https://doi.org/10.1155/2015/680381).
- [76] S. Dhyani, A. Kumar, and S. Choudhury, "Analysis of ECG-based arrhythmia detection system using machine learning," *MethodsX*, vol. 10, Mar. 2023, Art. no. 102195, doi: [10.1016/j.mex.2023.102195](https://doi.org/10.1016/j.mex.2023.102195).
- [77] S. Dixit and R. Kala, "Early detection of heart diseases using a low-cost compact ECG sensor," *Multimedia Tools Appl.*, vol. 80, nos. 21–23, pp. 32615–32637, Sep. 2021, doi: [10.1007/s11042-021-11083-9](https://doi.org/10.1007/s11042-021-11083-9).
- [78] N. N. Anuar, A. H. Hafifah, S. M. Zubir, A. Noraidatulakma, J. Rosmina, M. Y. N. Ain, H. M. Akma, Z. N. Farawahida, K. A. A. Shawani, M. A. D. Syakila, K. M. Arman, and A. J. Rahman, "Cardiovascular disease prediction from electrocardiogram by using machine learning," *Int. J. Online Biomed. Eng. (iJOE)*, vol. 16, no. 7, pp. 34–48, Jun. 2020, doi: [10.3991/ijoe.v16i07.13569](https://doi.org/10.3991/ijoe.v16i07.13569).
- [79] S. Karthiga and A. M. Abirami, "Deep learning convolutional neural network for ECG signal classification aggregated using IoT," *Comput. Syst. Sci. Eng.*, vol. 42, no. 3, pp. 851–866, 2022, doi: [10.32604/csse.2022.021935](https://doi.org/10.32604/csse.2022.021935).



MOHAMED AASHIQ received the B.Sc.Eng. degree in computer engineering and the M.Sc. degree in data science from the University of Peradeniya, Sri Lanka, in 2012 and 2021, respectively. He is currently pursuing the M.Sc. degree in computer and embedded systems engineering with Universiti Putra Malaysia (UPM), Malaysia. He is a Lecturer with South-Eastern University, Sri Lanka. He has authored scientific articles in indexed journals and IEEE conferences. His scientific research interests include biomedical signal processing, machine learning, the IoT, and data science.



SHAIFUL JAHARI HASHIM received the B.Eng. degree in the field of electronics and communication engineering from the University of Birmingham, U.K., in 1998, the M.Sc. degree from the National University of Malaysia, in 2003, and the Ph.D. degree from Cardiff University, U.K., in 2011. He is a Professor with the Department of Computer and Communication Systems Engineering, Faculty of Engineering and Head of Network Modelling Laboratory, Institute for Mathematical

Research (INSPEM), Universiti Putra Malaysia (UPM). His research interests are signal processing, the Internet of Things (IoT), wireless systems, cloud computing, network security, and 6G. He has contributed to more than 150 technical and research publications.



MARSYITA HANAFI received the B.Eng. degree in telecommunication engineering from the University of Malaya, Malaysia, in 2000, the M.Sc. degree from Universiti Putra Malaysia, in 2006, and the Ph.D. degree from Imperial College London, U.K., in 2012. She is currently associated with the Department of Computer and Communication Systems Engineering, Universiti Putra Malaysia, as a Senior Lecturer. She has published more than 70 scientific articles in several

peer-reviewed journals and at indexed conferences, two book chapters, and five intellectual properties. Her scientific research interests include artificial intelligence, image processing, and the Internet of Things system integration.



FAKHRUL ZAMAN ROKHANI (Member, IEEE) received the Ph.D. degree from the University of Minnesota, USA. He is currently an Associate Professor with Universiti Putra Malaysia and heads the smart system and system-on-chip (S3oC) research group. He was with Intel Penang Design Center and Huawei Technologies as a Visiting Professor, a Visiting Scholar with the ASIC and Systems State Key Laboratory, Fudan University, and a Visiting Professor with several other

universities. He has (co)-authored more than 130 peer-reviewed articles, one country policy article, ten books/book chapters, three magazine articles, and 25 intellectual properties. His current research interests include low-power/energy-efficient system-on-chip (SoC) design and automation, the IoT system integration, and sensor design. In 2023, he co-founded a startup company focusing on low-power IPs. He serves as the Vice President of Education and Communications for CAS Society, from 2023 to 2026, and the Steering Committee of IEEE Future Direction-Global Semiconductors at IEEE. On the publication front, he has been an Associate Editor of TCAS-I, an Associate Editor of the CAS Society Newsletter, a Guest Editor of TCAS-I and TCAS-II, and the Technical Program/Publication Chair/Track chairs/Embedded Workshop Chair of several IEEE CASS conferences.



AHMED FAEQ HUSSEIN (Member, IEEE) received the B.Sc. degree in electrical engineering science from Al-Mustansiriyah University, Iraq, in 1998, the M.Sc. degree in computer engineering from the University of Technology, Iraq, in 2004, and the Ph.D. degree from Universiti Putra Malaysia, in 2018. He is currently attached to the Department of Biomedical Engineering, College of Engineering, Al-Nahrain University, Iraq, where he currently is an Associate Professor.

He has authored more than 50 scientific papers in various peer-reviewed and indexed journals and conferences. His professional and scientific research interests include biomedical signal processing, cloud computing, and the Internet of Medical Things (IoMT). Prior joining to the academia, he was a Biomedical Engineering professional for more than 20 years.

...