STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **13**, February 2025, pp 547–567.
Published online in International Academic Press (www.IAPress.org)

Synergistic Approaches for Accurate Arrhythmia Prediction: A Hybrid AI Model Integrating Higuchi Dimensional Fractal, RR-intervals and Attention-based Convolutional Neural Network in ECG Signal Analysis

Nadia Berrahou *, Abdelmajid El Alami , Rachid El Alami , Hassan Qjidaa

Computer Science, Signals, Automation and Cognitivism Laboratory, Faculty of Science,
Sidi Mohamed Ben Abdellah-Fez University, Fez, Morocco

Abstract In recent years, numerous methods for detecting arrhythmias using a 12-lead ECG have emerged, with deep learning approaches notably demonstrating effectiveness and gaining widespread adoption. However, the classification of inter-patient ECG data for arrhythmia detection remains a significant challenge. Despite the increased utilization of deep learning methodologies, a noticeable gap persists in achieving optimal performance in inter-patient ECG classification. In this paper, we introduce a new method based on a 1D deep learning model that incorporates an attention mechanism into convolutional neural networks for arrhythmia detection. 1D-CNN layers automatically extract morphological characteristics from ECG data, providing an accurate technique for spatial feature extraction. Simultaneously, the attention mechanism enables the model to focus on crucial segments of a signal. To enhance temporal context, four RR-interval features are included, and the potential of the Higuchi Dimensional Fractal is explored as a method for extracting additional features from ECG signals. Consequently, the classification layers benefit from the combination of both temporal and deep features, contributing to the final arrhythmia classification. We validated the proposed method using the MIT-BIH arrhythmia dataset, employing an inter-patient paradigm for model training and validation. Additionally, to assess its generalization ability, we tested it on the INCART dataset. The proposed method attained an average accuracy of 98.75% for three classes and 97.96% for four classes on the MIT-BIH arrhythmia dataset. On the INCART dataset, it achieves an average accuracy of 98.12% for three classes. The experimental results indicate the superiority of this method in comparison to existing methods for recognizing arrhythmias. Thus, our method demonstrates enhanced generalization and potential effectiveness in identifying arrhythmias in real-world datasets characterized by class imbalances, showcasing its practical applicability.

Keywords Arrhythmia detection, Convolutional neural network, Attention mechanism, RR-intervals, Higuchi Dimensional Fractal, Electrocardiogram signals, MIT-BIH arrhythmia dataset, INCART dataset

DOI:10.19139/soic-2310-5070-2091

1. Introduction

Cardiac arrhythmias, also known as abnormal heart rhythms, manifest as heartbeats that are either excessively fast, slow, or irregular. These arrhythmias pose a threat to individuals' lives by impeding the heart's ability to pump an adequate volume of blood to vital organs. This has posed a significant global health challenge for numerous years, representing approximately 32% of total deaths, as indicated by a 2019 report from the World Health Organization (WHO) [1]. Consequently, the early detection and prompt treatment of arrhythmias are imperative for ensuring survival.

The electrocardiogram (ECG) is an essential tool for diagnosing arrhythmias, providing insights into the heart's rhythm, rate, and various aspects of cardiac electrical activity. However, ECG signals are inherently nonlinear and non-stationary, presenting challenges for manual interpretation. Manual analysis of ECG recordings can be laborious and susceptible to errors, especially when dealing with extended recordings necessary for identifying

*Correspondence to: Nadia Berrahou (Email: nadia.berrahou@usmba.ac.ma).

ISSN 2310-5070 (online) ISSN 2311-004X (print)

sporadic arrhythmias [2]. As a result, there is a pressing need for automated methods to aid clinicians in detecting arrhythmias from ECG data.

Categorizing heartbeats on an electrocardiogram (ECG) is essential for identifying arrhythmias. The Association for the Advancement of Medical Instrumentation (AAMI) has delineated five primary classes for heartbeats: Normal (N), Supraventricular Ectopics (S), Ventricular Ectopics (V), Fusions (F), and Unclassified (Q)[3]. Problematic arrhythmias are predominantly observed in supraventricular (S) and ventricular (V) beats. Ventricular (V) beats exhibit distinct morphologies, but they are less common, underscoring the significance of accurate identification. Conversely, normal (N) and supraventricular (S) beats often share similar morphologies, making differentiation based solely on local features challenging.

In clinical practice, the time interval between consecutive heartbeats (RR-interval) is crucial, especially for recognizing premature S beats with shorter RR-intervals [4]. However, variations in heart rhythm across different patients complicate the classification task. Moreover, the sporadic occurrence of supraventricular (S) beats can introduce bias into automated classification methods.

In developing an accurate diagnostic system for cardiac arrhythmias using ECG signals, the initial step involves extracting relevant features from these signals. Various methods have been proposed for this feature extraction, representing a critical stage in developing an accurate and effective diagnostic system for cardiac arrhythmias. Statistical features [5], and higher-order statistic features [6] are commonly employed for signal analysis. Recent research has expanded upon these by introducing additional features, including morphological features, and interval features [4, 7], wavelet entropy [8], auto-encoder [9], entropy features [10, 11], and fractal dimensions [12].

Many studies have utilized conventional signal processing and machine learning methods for ECG arrhythmia detection in classification tasks, such as K-Nearest Neighbor (KNN), Random Forest (RF), Support Vector Machines (SVM), Linear Discriminants (LD), and Artificial Neural networks (ANN) [13]. Despite the numerous classification methods proposed to enhance feature quality, certain drawbacks persist in these algorithms. The effectiveness of classification hinges greatly on the choice of features, a process that may be susceptible to subjective influences. Additionally, these methods frequently encounter overfitting issues [14]. Furthermore, various factors may cause performance discrepancies, including high noise levels, variations in ECG signals among patients, heterogeneous ECG sensor types, and disparities in arrhythmia prevalence across datasets.

In comparison to traditional methods, deep learning architectures have demonstrated significant advantages, particularly with the recent advancements that have introduced novel neural networks capable of automatically extracting and classifying features. Numerous ECG classification techniques based on deep learning models have emerged in the literature [15], demonstrating impressive performance. However, these models are not without challenges. A primary concern is overfitting, where a model becomes too closely aligned with the training data, leading to reduced performance when encountering new, unseen data. This highlights the importance of careful model design and the application of regularization techniques to ensure better generalization. Additionally, deep learning models typically require substantial computational resources. Larger models with more layers demand more processing time and memory, which can hinder their efficiency in long-term ECG monitoring on low-power or mobile devices. Another significant challenge is the issue of class imbalance in ECG datasets, where certain arrhythmia types are underrepresented compared to others. This imbalance can cause classification algorithms to favor the majority classes, resulting in decreased accuracy for the minority classes. To address this, various techniques such as oversampling, undersampling, cost-sensitive learning, ensemble methods, and specialized loss functions are employed to manage class imbalance and improve overall classification performance.

In the pursuit of higher classification accuracy for arrhythmias in clinical settings, attention mechanisms [16, 17] have proven valuable in effectively selecting and emphasizing relevant feature information. Despite these advancements, there is still significant potential for further improvements in classification accuracy.

Motivated by the findings presented in the referenced studies, we proposed a novel deep learning model that leverages the strengths of both Convolutional Neural Networks (CNN) and attention mechanisms. This model is designed to operate on electrocardiogram (ECG) data, autonomously extracting informative dependencies and combining them with RR-interval features and Higuchi Dimensional Fractal to enhance the precision of arrhythmia detection. The key contributions of this work can be outlined as follows:

1. Proposed a novel deep learning model for inter-patient paradigm classification by incorporating an attention mechanism into a Convolutional Neural Network (CNN), facilitating enhanced focus on critical information within ECG signals.

2. Utilized the Higuchi Dimensional Fractal (HDF) as a valuable tool, showcasing its effectiveness in capturing the intrinsic complexity of ECG signals. Anticipated the pivotal role of specific HDF values in accurately distinguishing between four types of arrhythmias.

3. Validated our proposed model using the publicly available MIT-BIH Arrhythmia dataset, and its generalization capabilities were assessed on the INCART dataset, demonstrating high performance in identifying four classes of arrhythmias compared to models reported in the state-of-the-art.

The structure of the remainder of this paper is as follows: The literature review and description of ECG signals are provided in Section 2. The methodology for detecting ECG arrhythmias, including data preprocessing and the proposed model, is detailed in Section 3. The results of the evaluation and experiments are discussed in Section 4, and the conclusion is provided in Section 5.

2. Related work

2.1. ECG Signal

An electrocardiogram (ECG) visually depicts the heart's electrical activity, providing insight into its functioning. This recording is taken from the body's surface and is a valuable tool for diagnosing cardiac ailments. Cardiologists analyze the ECG graph to detect irregular heart rhythms, changes in the structure of the heart's chambers, including the atria and ventricles, and conditions like ischemia, myocardial infarction, and other heart disorders involving changes in electrical conduction.

The ECG waveform comprises five waves: P, Q, R, S, and T [18]. Each of these waves corresponds to a distinct stage in the heart's cycle of polarization and depolarization. The P-wave signifies the depolarization of the atria, indicating an atrial event. The PR interval measures the duration between the onset of the P-wave and the commencement of the Q-wave, reflecting the time taken for electrical impulses to traverse between the atria and ventricles. The QRS complex designates the depolarization of the ventricles, indicating a ventricular event, while atrial repolarization also transpires during this phase. The T-wave signifies the repolarization of the ventricles. Finally, the RR interval signifies the temporal gap between two successive QRS complexes.

2.2. Literature review

Over the decades, substantial progress has been made globally in the field of ECG signal classification, utilizing a variety of methodologies ranging from traditional approaches to advanced machine learning and deep learning techniques.

Early methods for ECG signal classification relied on a sequential process involving three steps: signal preprocessing, feature extraction, and classifier design. The primary aim of preprocessing was to remove various forms of noise, such as motion artifacts and power line interference, from ECG recordings. Numerous denoising techniques have been explored, including low-pass filters, adaptive filters [19], and filter banks [20]. After preprocessing, important fiducial points were correctly identified and extracted, which provided useful morphological information about the heartbeats, such as the P wave, R peak, T wave, and QRS complex. Several algorithms and techniques were developed to precisely locate these fiducial points within the ECG signal [21, 22]. The next step involved extracting features from ECG segments. Many algorithms for classifying ECG beats relied on manually engineered feature extraction techniques. The advent of machine learning, particularly supervised classification techniques [23, 24, 25], transformed ECG signal classification. Algorithms such as Decision Trees (DT), K-nearest Neighbors (KNN), Support Vector machines (SVM), and Artificial Neural Networks (ANN [26, 27, 28, 29] have demonstrated considerable success in classifying ECG signals. Elhaj et al. [30] combined Principal Component Analysis (PCA) with Discrete Wavelet Transform (DWT) coefficients to classify ECG signals, achieving an impressive intra-patient classification accuracy of 98.91%. Ye et al. [26] employed a combination of Independent Component Analysis (ICA) and Discrete Wavelet Transform (DWT) to extract morphological features from ECG signals. These features were then integrated with dynamic RR-interval features and classified using a Support Vector Machine (SVM). This approach yielded an overall sensitivity of 53.46% and a precision of 62.79%. Zhang et al. [2] proposed a novel methodology for selecting disease-specific features to investigate their significance in distinguishing different types of heartbeats. Their study reveals that the RR interval emerges as the most effective feature in discerning between diseased and normal heartbeats. Additionally, they find that morphological distance, quantified by Dynamic Time Warping (DTW) distance between a given beat and the median beat in a recording, is particularly valuable for distinguishing between different types of diseased heartbeats. In this research, they employ features selected from both ECG leads and input them into a combined Support Vector Machine for the classification of heartbeats. Their model attains an inter-patient classification overall accuracy of 86%. Shi et al. [31] proposed a hierarchical classification approach employing weighted extreme gradient boosting (XGBoost) and recursive feature reduction, resulting in an inter-patient classification accuracy of 92.1%. Chen et al. [32] integrated weighted RR-interval features with projected ECG features, demonstrating high classification performance in the intra-patient evaluation paradigm but encountering reduced sensitivity and precision in the inter-patient evaluation paradigm.

Although these methods achieved high classification rates, they heavily relied on the quality of features extracted from ECG signals. Many frameworks treated ECG signals as sequences of stochastic patterns, necessitating complex feature extraction processes, and high sampling rates. However, the robustness of these techniques was limited due to significant intra-class variation observed in ECG signals. The emergence of deep learning marked a significant turning point in ECG signal classification. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) demonstrated their ability to autonomously extract discriminative features. CNNs, in particular, gained widespread adoption in ECG classification due to their capability to automatically capture crucial patterns and features.

Within the intra-patient paradigm of ECG beat classification, various approaches were proposed. For instance, Acharya et al. [15] introduced a nine-layer Convolutional Neural Network (CNN) to classify five categories of ECG beats using the MIT-BIH arrhythmia dataset. They addressed the class imbalance challenge among these categories using a data augmentation technique, resulting in an impressive overall classification accuracy of 94.03%. Other approaches for heartbeat classification, including CNN-based models, were proposed in references [33, 34]. Ubeyli [35] introduced a classifier based on Recurrent Neural Networks (RNN) with an eigenvector method, achieving an average accuracy of 98.06% in detecting four distinct arrhythmia classes. Shaker et al. [36] introduced a data augmentation technique utilizing GANs to address the class imbalance issue in the MIT-BIH arrhythmia dataset. They employed two distinct deep learning approaches, both relying on deep CNNs. The first approach achieved an impressive overall accuracy rate of 98.30% and a precision of 90.00%, while the second approach attained an overall accuracy of 98.00% and a notably higher precision of 93.95%. Moreover, Long Short-Term Memory Networks (LSTMs) have emerged as a powerful tool for discerning between various beat classes. Studies cited in [37, 38] developed arrhythmia detection models based on LSTMs, showcasing robust recognition performance across five different beat categories while mitigating computational overhead. This highlights the efficacy of LSTMs in capturing temporal dependencies in ECG data. Additionally, a hybrid model that merges Convolutional Neural Networks (CNNs) and LSTMs was introduced in [39] for arrhythmia detection. This hybrid approach harnesses both temporal and spatial information in ECG signals. Particularly noteworthy is its utilization of variable-length ECG segments, enhancing its adaptability in the context of arrhythmia detection. Within the context of inter-patient beat classification, various approaches were developed to classify heartbeats from different patients. For example, Sellami et al. [40] devised a robust deep Convolutional Neural Network for arrhythmia classification without denoising the ECG signals. Subsequently, Zhang et al [41] proposed a CNN-based adversarial deep learning model for the inter-patient evaluation paradigm, comprising an encoder, classifier, and adversary networks. The classifier integrates input features such as ECG segments and normalized local and global RR intervals, treated as distinct channels. While the model exhibited promising results, its complexity may limit its applicability, particularly in edge-computing scenarios. Alternatively, T.Wang et al. [42] presented an inter-patient ECG model employing CNNs and Continuous Wavelet Transform (CWT). A 2D CNN was then employed alongside CWT-generated time-frequency scalograms of ECG segments and RR intervals for beat classification. Nevertheless, the CWT preprocessing step introduced additional computational overhead to the classifier, potentially restricting its suitability for edge inference scenarios. Shi et al. [43] presented a multi-input neural network architecture that leveraged RR interval features and divided the heartbeat into three segments as input. Their model achieved an impressive accuracy of 94.2% for four classes in the inter-patient paradigm. Niu et al. [44] proposed an innovative deep learning approach for the inter-patient paradigm, incorporating symbolization techniques and a multi-perspective Convolutional Neural Network. Their method achieved a high accuracy of 96.4%. Li et al. [45] introduced a 3-D input format and employed a multi-field CNN to attain notable accuracy on public datasets. Additionally, Guo et al. [46] combined Dense CNN and Recurrent Neural Network architectures to process long-term ECG segments, resulting in impressive performance in arrhythmia detection.

Despite these notable research efforts, several challenges persist in arrhythmia detection. These challenges primarily revolve around two key issues. Firstly, the implementation of deep learning techniques often necessitates high-performance computing resources, posing difficulties in resource-constrained environments. Secondly, these techniques may exhibit bias stemming from the characteristics of the training data, resulting in reduced accuracy when confronted with new or diverse datasets. The proposed study introduces a novel ECG classification model employing an Attention-based Convolutional Neural Network to address these limitations. Evaluated within the inter-patient paradigm, this model demonstrates superior performance compared to existing literature. By leveraging the advantages of deep learning, this approach enhances the accuracy and effectiveness of ECG signal classification.

3. Materials and Methods

The proposed method comprises three essential components: preprocessing, feature extraction, and classification, as visually depicted in Figure 1. The preprocessing phase incorporates two key steps: baseline wander removal and heartbeat segmentation. The second phase is dedicated to feature extraction, where the objective is to represent the changes in the original signal precisely. Finally, the third phase involves the classification of arrhythmias. Below, we provide an overview of each of these components.

3.1. Description of datasets

In this study, two distinct ECG datasets are used: the MIT-BIH-arrhythmia dataset and the INCART dataset, both of which are publicly accessible through PhysioBank [47]. While most research on heartbeat classification has historically utilized the MIT-BIH arrhythmia dataset for training and evaluation [48], our model undergoes training and evaluation exclusively on this dataset to ensure fair comparisons. Additionally, we utilize the INCART dataset to evaluate the generalization ability of the proposed model.

All ECG recordings within these datasets share a uniform duration of 30 minutes. However, they differ in sampling frequency. Before utilization, it is necessary to resample the recordings to a uniform 360 Hz frequency. The recordings come equipped with comprehensive heartbeat-level annotations. These annotations initially consist of 15 distinct classes, which are later grouped into 5 superclasses based on the AAMI [3]. Among them, class Q was excluded from the analysis due to the limited number of samples available. Further details regarding these datasets are provided below.

- MIT-BIH arrhythmia dataset

  The dataset comprises 48 ambulatory ECG recordings with two leads from 47 patients, each identified by a 3-digit number. The recordings were sampled at 360 Hz per channel with 11-bit resolution over a 10 mV range. Typically, the first lead used is the modified limb lead II (MLII), except for recording 114, where a different lead is utilized. However, for this study, only the MLII lead is utilized. The recordings containing paced beats (i.e., 102, 104, 107, and 217) were excluded per the (AAMI).

  As previously indicated, the primary goal of this paper is to evaluate the efficacy of the proposed method within the inter-patient paradigm. To enable a comparison with prior studies, we adopted a widely used

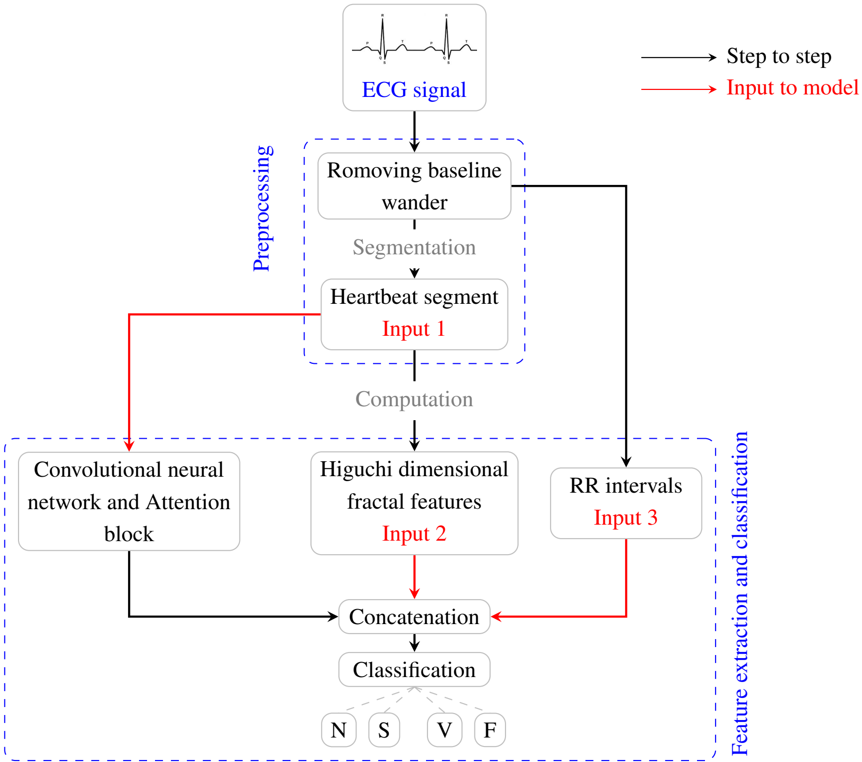

Figure 1. Proposed work-flow diagram for Arrhythmia classification.

data partitioning approach introduced by de Chazal et al. [4] for the MIT-BIH arrhythmia dataset. Each dataset, DS1 and DS2, comprises 22 records meticulously selected to maintain a similar distribution of different beat types. DS1 is designated for training purposes, while DS2 is reserved for evaluating the method's performance. Importantly, no patient overlaps between the two datasets. The division details and corresponding distribution of heartbeat classes are presented in Table 1.

- INCART 12-Leads arrhythmia dataset

  This dataset contains 75 ECG recordings, with a sampling frequency of 257 Hz, each comprising the standard set of 12 leads. In alignment with this study's methodology, only the modified limb lead II is utilized. The annotations for these recordings were initially generated using an automated algorithm and subsequently refined manually, adhering to the established PhysioBank beat annotation definitions. Notably, none of the recordings include pacemakers; however, most feature ventricular ectopic beats. The distribution of heartbeats within the INCART dataset is presented in Table 2.

3.2. Preprocessing

Preprocessing the original data is crucial as an initial phase to achieve optimal performance of the proposed model. This phase primarily consists of the following two steps.

3.2.1. Removing baseline wander The interference caused by noise artifacts in the ECG signal, referred to as Baseline Wander, typically falls within the frequency range of 0.15 to 0.3 Hz and often originates from factors such as breathing apparatus. Eliminating Baseline Wander is essential for enhancing the accuracy of ECG analysis, particularly in reducing irregularities associated with heartbeats. Various methods can be employed to eliminate or

Table 1. The inter-patient paradigm distribution in MIT-BIH arrhythmia dataset.

Datasets	N	S	V	F	Records
DS1	45866	944	3788	415	101, 106, 108, 109, 112, 114 115, 116, 118, 119, 122, 124 201, 203, 205, 207, 208, 209 215, 220, 223, 230
DS2	44258	1837	3221	388	100, 103, 105, 111, 113, 117 121, 123, 200, 202, 210, 212 213, 214, 219, 221, 222, 228 231, 232, 233, 234

Table 2. Heartbeat distributions in the INCART dataset.

Dataset	N	S	V	F
INCART	153491	1958	19993	219

mitigate Baseline Wander from the ECG signal. In this study, two consecutive filters are utilized [49]: a median filter with a width of 200 ms followed by a median filter with a width of 600 ms. Subsequently, the resulting baseline is obtained by subtracting it from the raw ECG signal. (see Figure 2).

3.2.2. Data segmentation Before undertaking ECG classification, a critical preliminary step involves the extraction of individual heartbeats from the ECG signal. This process typically involves accurately detecting QRS waves and identifying fiducial points within each heartbeat. In our specific study, we achieved this task by segmenting the ECG signal into distinct heartbeats. This segmentation was accomplished by utilizing the positional information of R-peaks labeled in the MIT-BIH arrhythmia dataset. Each extracted heartbeat consisted of 200 sample points, meticulously centered around the position of the R-peak, as illustrated in Figure 3.

3.3. Features extraction

3.3.1. RR-Interval The RR interval, which represents the time duration between consecutive R-peaks on the ECG waveform, plays a crucial role in heartbeat classification. Empirical evidence, as detailed in [2], underscores the significance of RR intervals in differentiating between various heartbeat conditions, particularly in distinguishing normal beats from abnormal conditions like supraventricular premature beats and ectopic beats. This study introduces the extraction of four specific types of RR intervals from ECG signals, as described in [50], the previous RR-interval, post-RR-interval, local RR-interval, and the ratio RR-interval.

The computational complexity of extracting four specific types of RR intervals depends on the algorithms and techniques used for extraction. This complexity is influenced by the computation of interval durations. Each type of RR interval extraction typically involves linear complexity, as it requires basic arithmetic operations or averaging over detected R-peaks.

3.3.2. Higuchi Dimensional Fractal (HDF) Fractals can be comprehended through either statistical or morphological interpretations of self-similarity within an object. Self-similarity entails the object reproducing its patterns across multiple scales, preserving its scale invariance. This self-replicating property holds significant importance as it establishes enduring connections between measurements conducted at diverse scales and resolutions. This distinctive trait is indispensable for gaining insights into an object's behavior across a broad spectrum of scales. The Fractal Dimension (FD) is introduced as a metric to gauge the extent of self-similarity and the scaling characteristics of an object or dataset. It provides a means to depict the intricacy and level of detail an object possesses as you zoom in or out. Diverse methods are available for calculating the FD, each rooted in various

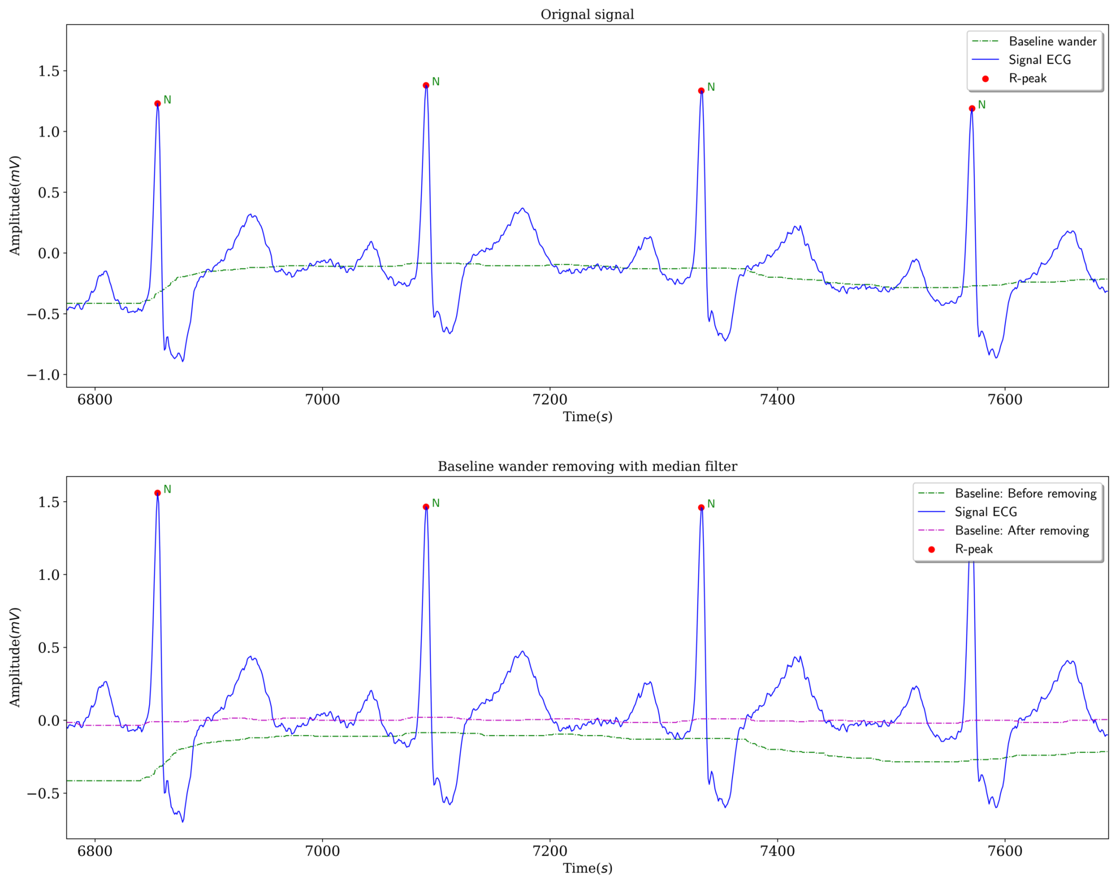

Figure 2. The ECG signal before and after applying the median filter.

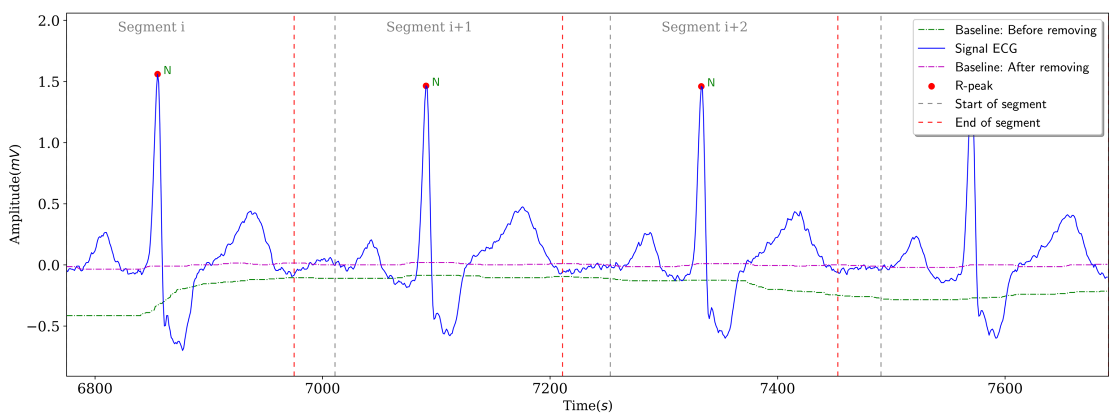

Figure 3. Segmentation of ECG Signal.

polyscalar relationships, such as variance, length, correlation, entropy, and spectral properties. In the context of this work, our primary focus will be on the Higuchi Dimensional Fractal (HDF).

The Higuchi Dimensional Fractal (HDF) serves as a nonlinear measure of waveform complexity when applied in the time domain [51]. For a discretized signal represented as a sequence of data points $X = \{X(1), X(2), \ldots, X(N)\}$, where N is the total number of samples, the Fractal Dimension is computed as follows:

First, a set of k new self-similar time series X_m^k is constructed as:

$$X_m^k = \left\{ X(m), X(m+k), X(m+2k), X\left(m + \left[\frac{N-k}{k}\right] k\right) \right\} \tag{1}$$

for $m = 1, 2, ..., k$ where m represents the initial time, k denotes the discrete time interval between points ($k = 1, 2, \ldots, k_{max}$) with k_{max} being a free parameter and [.] denoting the floor function.

The length, $L_m(k)$, of each curve X_m^k is then determined as follows :

$$L_m(k) = \frac{N-1}{\left[\frac{N-m}{k}\right] k^2} \sum_{i=1}^{\left[\frac{N-m}{k}\right]} |X(m+ik) - X(m+(i-1)k)| \tag{2}$$

where $\frac{N-1}{\left[\frac{N-m}{k}\right] k}$ is a normalization factor.

Next, the mean value of the curve length $L(k)$ for each $k = 1, 2, \ldots, k_{max}$ is obtained by averaging $L_m(k)$ for all m, expressed as:

$$L(k) = \frac{1}{k} \sum_{m=1}^{k} L_m(k) \tag{3}$$

The relationship between $L(k)$ and k follows a power law, represented as $L(k) \propto k^{-\text{HDF}}$. To estimate the Higuchi Dimensional Fractal (HDF) for an ECG signal, the slope of the linear least-squares fit to the plotted points $\{ln(L(k)), ln(1/k)\}$ is calculated. This slope provides the fractal dimension HDF, determined using the formula for each $k = 1, 2, \ldots, k_{max}$:

$$HDF = \frac{\ln(L(k))}{\ln(1/k)} \tag{4}$$

The values of the Higuchi Dimensional Fractal (HDF) were obtained by varying the time interval parameter, k, within a specific range in the study. Using different values of k within this range generated a corresponding series of HDF values, which were utilized as a feature in the analysis.

Selecting the parameter k_{max} is crucial for obtaining an accurate estimation of the Higuchi Dimensional Fractal (HDF). This parameter defines the maximum number of new self-similar time series generated from the original signal. A k_{max} that is too low may not capture the full complexity of the signal, while a k_{max} that is too high can introduce noise and make the estimation unstable. To determine the optimal k_{max}, it is common to plot the HDF as a function of different k_{max} values and identify local maxima or plateaus, where the HDF estimation becomes relatively stable. The optimal k_{max} is typically the one that corresponds to these points, as it represents a good balance between accuracy and noise sensitivity. In practice, it is recommended to choose the smallest k_{max} that provides a reliable estimate, in order to minimize computational complexity while maintaining the precision of the analysis. In the computations detailed in the results section, a maximum value of k_{max}, was set to 30, chosen for its effectiveness in achieving the desired analytical results and accuracy.

The computational complexity of calculating the HDF is dominated by the construction of self-similar time series and the calculation of curve lengths, which have a linear complexity for the number of samples in the signal.

Therefore, the total computational complexity of the HDF calculation is $O(k_{max} \times N)$, where N is the length of the input signal.

3.4. Description of the proposed model

3.4.1. Overview and motivation Detecting irregular ECG signals relies on two key aspects: identifying abnormalities in the shape of the ECG waveform and spotting deviations in the timing of irregular ECG fluctuations. A model designed for this purpose must excel in feature extraction to precisely identify various ECG patterns. Convolutional Neural Networks have gained traction for their exceptional performance in various tasks, particularly in domains such as image and speech recognition, thanks to their ability to extract local features from input data autonomously.

However, it's crucial to acknowledge that a standalone CNN, originally intended for data classification, might not inherently capture the crucial temporal aspects required for ECG analysis. To address this limitation, combining CNNs with RR intervals presents itself as a logical approach. This combination effectively manages temporal information while leveraging the CNN's feature extraction capabilities. Moreover, incorporating fractal dimensions enhances the model's ability to characterize the intricacies within the ECG signal. Additionally, the introduction of an attention mechanism is expected to bring benefits in two aspects. Firstly, it could assist the model in directing its attention towards critical information within ECG signals, enhancing classification accuracy. Secondly, it could assist in identifying abnormal signal patterns and their specific locations, thereby enhancing the model's interpretability.

3.4.2. Structure of the proposed model This study proposes a new method, depicted by the model structure in Figure 4. In this model, the ECG signal undergoes initial processing through the segmentation module, yielding several segmented beats. These beats are then directed to the feature extraction module, which extracts features for each beat using convolutional blocks. Subsequently, these features enter the inter-beat feature fusion module, where an attention block is employed. This attention block assigns varying weights to accentuate valuable beats while attenuating less informative ones. The resultant fused features are then concatenated with manually extracted RR-interval values and HDF features. Finally, two fully connected layers act as a classifier, producing the classification probability.

- Convolutional block (BlockConv)

  The proposed model comprises three convolution blocks, each consisting of a convolution layer, an instance normalization (IN) layer, and the Mish activation function. The model's input is preprocessed electrocardiogram data in one-dimensional form.

  In the first block, a single convolution layer with a kernel size of 8 and 16 filters is utilized. Subsequently, in the following two blocks, the convolution kernel size is reduced to 6, with filter numbers of 32 and 48, respectively.

  An adaptive max pooling layer is applied to reduce the dimensionality of the feature map. Additionally, a flatten layer is used to convert the output from the previous layer into a single vector.

- Attention block

  The attention block, as illustrated in Figure 4, is pivotal in our model for effectively incorporating the self-attention mechanism to combine features extracted from different heartbeats. In the context of electrocardiogram (ECG) signal analysis, where capturing intricate patterns and dependencies is crucial, the self-attention mechanism plays a vital role. ECG signals often exhibit dynamic variations and irregularities, and the self-attention mechanism enables our model to selectively prioritize certain segments or patterns within the signal, allowing for a more nuanced understanding of the underlying cardiac activity. This adaptability is particularly significant in ECG analysis, as it empowers the model to discern subtle anomalies or variations that might be indicative of cardiac abnormalities. Therefore, the integration of the self-attention mechanism within the attention block enhances the model's ability to discern and leverage crucial information from diverse beats within the ECG signal.

The attention mechanism as defined in reference [52] operates through the following steps: starting with a set of n feature vectors denoted as $F = [f_1, f_2, \ldots, f_n]$. Each vector f_i is first transformed using the non-linear activation function $\tanh(.)$, producing a set of activated vectors $U = [u_1, u_2, \ldots, u_n]$. Each activated vector u_i is then processed through a series of k parallel linear layers, resulting in k different representations for each vector u_i. These representations are mathematically expressed as:

$$u_i = \tanh(W_i f_i + b_i), \tag{5}$$

$$v_{i,j} = W_{i,j} u_i + b_{i,j}, \quad \text{for} \quad j = 1, 2, \ldots, k \tag{6}$$

where $W_{i,j}$ and $b_{i,j}$ are the weight matrices and biases for the j-th linear layer applied to u_i.

After obtaining the k representations for each u_i, the softmax function is applied to these representations to calculate attention weights. The softmax function normalizes these values to a range between 0 and 1, providing the relative importance of each representation:

$$a_{i,j} = \frac{\exp(v_{i,j})}{\sum_{l=1}^{k} \exp(v_{i,l})}, \tag{7}$$

where $a_{i,j}$ represents the attention weight for the j-th layer of the vector u_i.

Finally, a weighted sum of the k representations is computed using the attention weights to produce the attention output vector $f_{\text{att},i}$ for each input feature vector f_i:

$$f_{\text{att},i} = \sum_{j=1}^{k} a_{i,j} v_{i,j} \tag{8}$$

This attention block enhances the ability of the model to focus on the most informative features by leveraging multiple parallel linear layers, enabling it to capture complex relationships between input features and improve performance on tasks such as classification and signal analysis.

4. Experiments and Results

4.1. Evaluation metrics

In this paper, we assessed the performance of our model using three metrics, including overall accuracy (Acc), sensitivity (Sen), and precision (Ppr) for each class. These metrics are calculated as follows:

$$Acc = \frac{T_p + T_n}{T_p + F_p + F_n + T_n} \tag{9}$$

$$Sen = \frac{T_p}{T_p + F_n} \tag{10}$$

$$Ppr = \frac{T_p}{T_p + F_p} \tag{11}$$

where T_p represents true positives, T_n indicates true negatives, F_n stands for false negatives, and F_p signifies false positives.

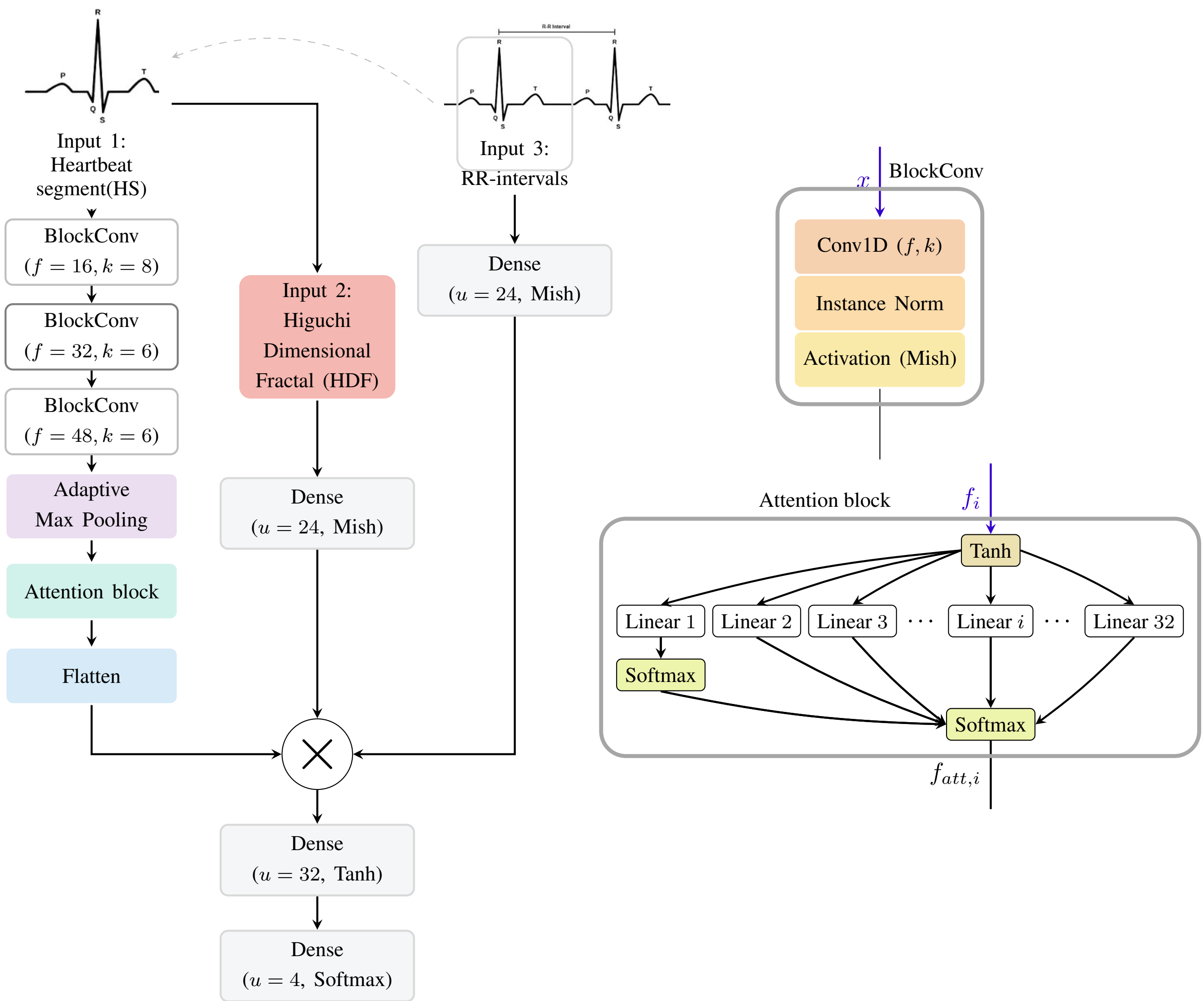

Figure 4. Architecture of the proposed model.

4.2. Training process

The raw dataset is split into a training dataset and a testing dataset. The training process consisted of 50 epochs, with a batch size of 512 for the model. We utilized the Rectified Adam algorithm with a weight decay of 0.0034 for model optimization. Additionally, a learning rate decay strategy is used, reducing the learning rate every 7 epochs with a decay factor gamma of 0.65. Initially, the learning rate is set to 0.0275. This gradual reduction in the learning rate helps stabilize the model in later epochs. We employed the cross-entropy algorithm as the loss function, incorporating weighted loss due to the class imbalance present in the dataset. Our model is trained on a system equipped with a single NVIDIA V100 GPU boasting 32 GB of RAM.

4.3. Results and Comparison

The proposed heartbeat detection model's performance is assessed through the inter-patient evaluation method using an experimental dataset sourced from MIT-BIH. This dataset is widely employed in ECG research due to its comprehensive expert annotations, as described in Section 3. Our experimentation is structured into two distinct parts, each designed to assess the effectiveness and generalization capability of our model. In the first part, we adhere to a common research practice by training the model on the DS1 subset of the MIT-BIH arrhythmia dataset and then evaluating its performance on the DS2 subset. In the second part of our experiments, we focus on cross-dataset evaluations. For this purpose, we utilize both DS1 and DS2 subsets from the MIT-BIH arrhythmia dataset to train and validate our model. Once trained, the resulting model is tested using the INCART dataset.

Table 3. Comparison results of our proposed method with existing works for the inter-patient paradigm.

Model	Class	Acc (%)	N		S		V		F	
			Ppr(%)	Sen(%)	Ppr(%)	Sen(%)	Ppr(%)	Sen(%)	Ppr(%)	Sen(%)
Sellami et al. [40]	3	95.08	-	-	-	-	-	-	-	-
Zhang et al. [41]	3	94.70	98.00	96.20	90.80	78.80	94.30	92.50	-	-
Garcia et al. [53]	3	92.4	98.00	94.00	53.00	62.00	59.40	87.30	-	-
Lin and Yang [7]	3	93	**99.30**	91.60	31.60	81.40	73.70	86.20	-	-
T.Wang et al. [42]	3	97.75	98.65	99.25	80.37	65.58	93.46	95.46	-	-
Zahid et al. [54]	3	98.19	98.85	99.30	83.51	**83.37**	**97.38**	91.39	-	-
Farag [55]	3	98.18	99.00	99.10	82.68	81.60	95.63	95.00	-	-
Xu et al. [56]	3	98.50	98.64	98.87	82.48	83.06	90.04	93.46	-	-
Berrahou et al. [57]	3	98.73	99.05	**99.73**	**95.02**	79.02	96.79	96.33	-	-
HS+RR	3	98.32	98.84	99.57	90.96	75.69	94.65	94.00	-	-
HS+RR+HDF	3	**98.75**	99.01	**99.73**	94.58	78.91	97.22	**96.68**	-	-
Shi et al. [31]	4	92.1	**99.5**	92.1	46.2	**91.7**	88.1	95.1	15.12	61.6
Shi et al. [43]	4	94.2	98.92	95.26	47.85	90.74	84.49	92.92	-	-
Niu et al. [44]	4	96.4	97.4	98.9	76.6	76.5	94.1	85.7	1.79	0.25
Li et al. [45]	4	91.44	98.92	91.81	35.41	89.05	90.11	95.15	20.36	32.22
G.Wang et al. [58]	4	94.62	98.13	96.06	82.89	71.92	80.16	92.01	23.32	**62.64**
Xia et al. [59]	4	94.69	97.41	97.44	56.73	68.41	88.70	82.91	-	-
Zhou et al. [60]	4	95.6	98.8	96.9	53.8	89.3	92.3	93.3	-	-
Chen et al. [61]	4	96.77	97.81	99.15	76.10	63.90	**96.78**	88.84	**54.25**	47.68
HS+RR	4	97.76	98.23	**99.55**	**91.00**	78.75	95.10	95.78	00.00	00.00
HS+RR+HDF	4	**97.96**	98.46	99.54	90.28	82.51	95.89	**97.05**	3.57	0.26

4.3.1. Inter-Patient Classification This section focuses on validating our method using the inter-patient paradigm for heartbeat classifiers. Table 3 displays the results of inter-patient classification involving three and four classes using precision and sensitivity metrics for each class, along with the average accuracy computed for the entire dataset. The models compared in Table 3 include the HS+RR model and its extension with HDF features, denoted as HS+RR+HDF.

Among the models evaluated for classifying arrhythmias into three classes, HS+RR exhibits notable accuracy at 98.32%, with corresponding sensitivity rates for classes N, S, and V standing at 99.57%, 75.69%, and 94.00% respectively. Our model, HS+RR+HDF, achieves an outstanding average accuracy of 98.75% for three classes. It also demonstrates sensitivity scores of 78.91% and 96.68% for S and V, respectively.

On the other hand, the HS+RR model achieves an accuracy of 97.76%, with respective sensitivities of 99.55% for class N, 78.75% for class S, and 95.78% for class V. However, its sensitivity for the minority class F is 0.00%, indicating potential challenges in classifying this class. Incorporating HDF features in HS+RR+HDF leads to a significant improvement in sensitivity for class F, resulting in an increased accuracy of 97.96% for four classes. It shows lower average scores due to challenges in F class detection due to the limited number of class F heartbeats and their resemblance to class N. However, our models maintain satisfactory sensitivity in the N, S, and V classes.

These results lead to several observations. Our models presented in Table 3 demonstrate high performance when faced with the challenge of inter-patient ECG classification. Notably, in the overall classification performance, class V demonstrates superior performance compared to class S. This variance might be influenced, in part, by the smaller sample size of class S but its greater number of subclasses compared to class V.

The HS+RR+HDF model consistently surpasses HS+RR across all metrics for N, S, and V in the three-class classification, as well as sensitivity for S and V in the four-class classification. This indicates that integrating HDF as a feature significantly enhances the model's performance. As a result, it enhances precision and sensitivity in distinguishing between different cardiac arrhythmias, thereby facilitating more accurate diagnoses.

Figure 5 and Figure 6 present a comparison of accuracy and loss between two models, HS+RR+HDF and HS+RR, during the training process using the test set for the inter-patient paradigm with three and four classes respectively. The HS+RR model exhibits oscillation during training, whereas the HS+RR+HDF model demonstrates superior stability. As training progresses, accuracy gradually improves and eventually stabilizes. Additionally, during epochs 20 to 50, the model consistently attains reduced and stabilized loss values across the training and validation datasets. The proposed model, integrating RR intervals and HDF features, shows notable stability and accuracy. Moreover, it achieves faster convergence and avoids overfitting compared to the HS+RR model.

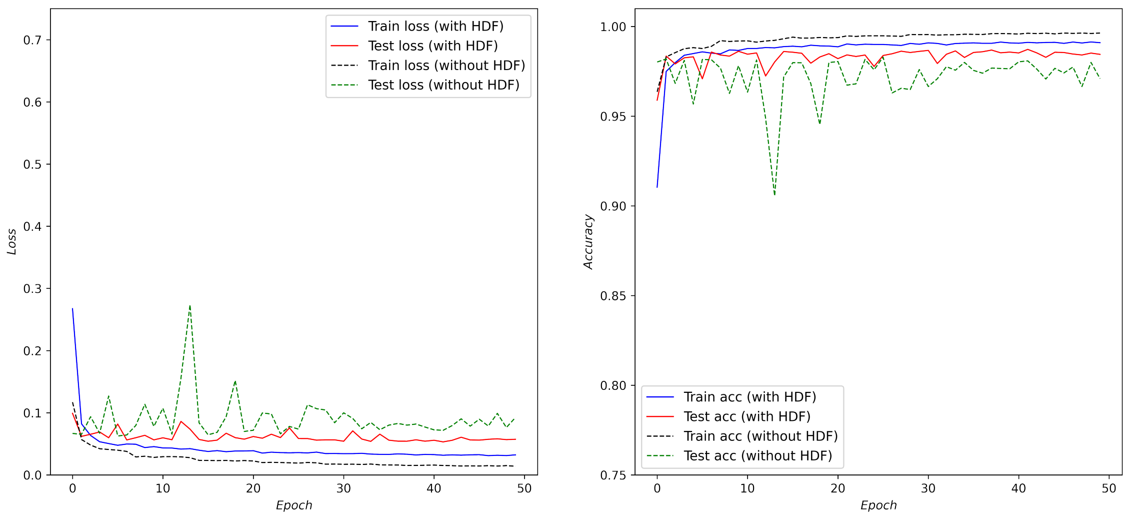

Figure 5. Training/test accuracy and loss curves for the inter-patient paradigm (3 classes).

Our proposed model is compared with existing models using the identical test set for three and four classes. Table 3 illustrates the summarized comparative results. The adversarial CNN model described by [41] achieved sensitivities of 78.8% for S and 90.8% for V, with corresponding precisions of 92.5% and 94.3%, respectively. Zahid et al. [54] introduced a 1D Self-Organizing Neural Network (ONN) for ECG arrhythmia classification, achieving sensitivities of 83.37% for S and 91.39% for V. Xu et al. [56] proposed an model for arrhythmia classification by incorporating residual and attention mechanisms in the U-Net. achieved an accuracy of 98.5% and sensitivities for the classes S and V of 83.06% and 93.46%, respectively. Berrahou et al. [57] proposed a new method for arrhythmia detection in inter patient by using RR intervals and entropy rate with CNN architecture, achieving sensitivities of 79.02% for S and 96.33% for V. While methods presented in [41, 54, 56, 57] exhibit some advantages, particularly for the S class, our proposed method surpasses these methods in terms of average accuracy and excels in all metrics for other classes. Farag [55] presented a CNN operating on the first derivative of the ECG signal, leveraging matched filters to reproduce different arrhythmia classes within convolutional layers. Additionally, the model processed RR intervals through a stack of fully connected layers, achieving an accuracy of 98.18%. Lin and Yang [7] introduced a method based on normalized and non-normalized RR intervals, extracting ECG morphology through wavelet analysis and a linear prediction model. However, this method exhibits a lower

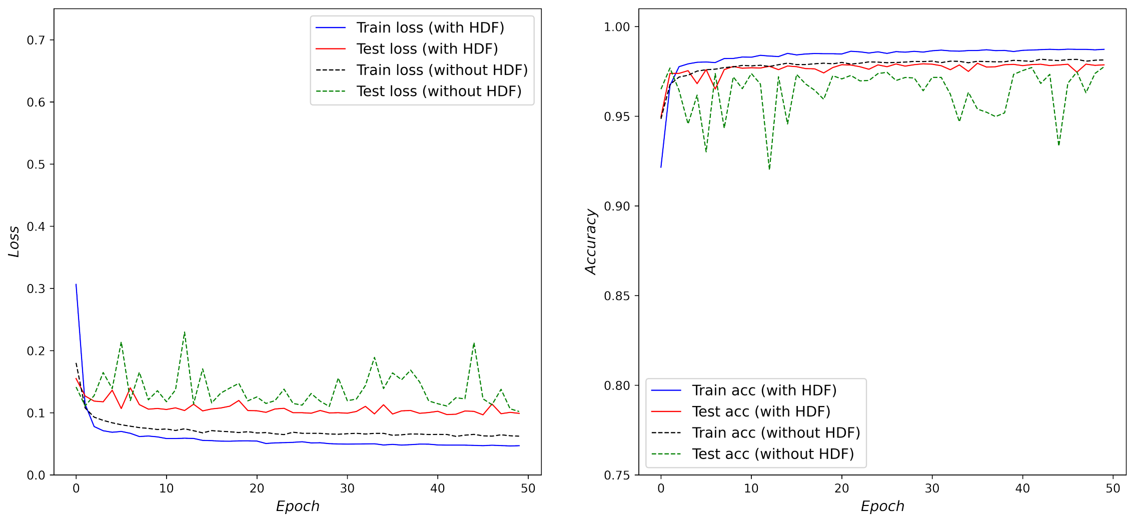

Figure 6. Training/test accuracy and loss curves for the inter-patient paradigm (4 classes).

prediction accuracy than our proposed method. Ultimately, our method surpasses those proposed by [40, 42, 53] in all aspects, as shown in Table 3.

Regarding four class classifications, our method demonstrates superior performance compared to all techniques with four classes listed in table 3 in terms of accuracy. Niu et al. [44] introduced an innovative approach for inter-patient heartbeat classification, utilizing a symbolization method and a multi-perspective convolutional neural network. While achieving a high classification accuracy of 96.4%, their model exhibited a low sensitivity score for class F (0.25%). Shi et al. [31] developed a pair of hierarchical weighted XGBoost classifiers, achieving impressive sensitivity values for N, S, and V of 92.1%, 91.7%, and 95.1%, respectively. Li et al. [45] proposed a DNN-based automatic heartbeat classifier involving a complex process for beat-to-beat correlation calculation. Although methods presented in [31, 45] demonstrated some advantages over our approach, particularly for the S class, the overall performance fell short of expectations. Shi et al. [43] introduced a deep neural network with multiple inputs for classifying heartbeats, achieving the best sensitivity result for S beats at 91.7%, yet the precision for the S class was only 47.85%, with an overall accuracy of 94.2%. Meanwhile, Xia et al. [59] suggested a novel generative adversarial network to enhance minority class features, yielding sensitivity values of 97.44%, 68.41%, and 82.91% for N, S, and V, respectively. Despite these advancements, the overall classification performance was inferior to our proposed method, primarily attributed to inherent limitations in generative models. Additionally, G.Wang et al. [58] utilized CNN to construct an ECG classification model in an unsupervised manner, achieving a sensitivity of 92.01% for the V class, however, the sensitivity for the S class was only 71.92%. Compared to Chen et al. [61], which demonstrates the best performance reported, our method shows an increase in accuracy by 1.19%, as well as an improvement in the sensitivity for the S and V classes by 18.61% and 6.21% respectively. However, it does lead to a decrease in the sensitivity for the F class by 47.42%. The clear trend in Table 3 is that most of the listed studies faced challenges detecting S and F heartbeats. In summary, this comparative analysis stands out ECG heartbeat classification results with the proposed method within the inter-patient paradigm. These results collectively underscore the effectiveness of our approach in accurately identifying S and V arrhythmia cases. Nevertheless, there is still potential for enhancing the performance of the F class.

4.3.2. Generalization of the proposed model The model is trained and validated during this experiment using DS1 and DS2 datasets, as detailed in Table 1. Following this, the trained model underwent testing on the INCART dataset. The summarized outcomes of the cross-dataset inter-patient validation experiment, focusing on three classes, are detailed in Table 4. Although the model's accuracy slightly decreased by 0.63% in this testing dataset, its performance in classifying N (normal) and V (ventricular ectopic beat) instances remained consistent. However, there was a noticeable decline in the classification of S (supraventricular ectopic beat) instances, with precision and sensitivity values of 69.40% and 59.65%, respectively. This drop in performance is attributed to variations in the training and testing datasets, acquired under different conditions and using distinct equipment. Nonetheless, the overall performance of our proposed classifier remains acceptable. Figure 7 shows the accuracy and loss curves for the HS+RR+HDF model trained on the DS1 dataset and tested on the Incart dataset in an inter-patient paradigm across three classes over 50 epochs. The curves indicate better convergence and no signs of overfitting. The proposed model, which integrates RR intervals and HDF features, demonstrates notable stability and accuracy. Importantly, our model is comparable to or better than other models listed in Table 4 concerning the overall accuracy classification and demonstrates comparable results for N, V, and S classes. This highlights the model's ability to generalize well to unseen data.

Table 4. Generalization performances (%) on INCART dataset.

Model	Test Acc (%)	N		S		V	
		Ppr(%)	Sen(%)	Ppr(%)	Sen(%)	Ppr(%)	Sen(%)
M Llamedo [62]	91.00	99.00	92.00	11.0	**85.00**	88.00	82.00
He et al. [63]	91.60	**99.60**	92.00	14.40	81.00	81.90	**91.00**
G.Wang et al. [58]	95.36	-	-	37.38	44.07	92.08	77.02
T.Wang et al. [42]	92.70	93.35	**99.70**	40.57	59.65	**98.34**	41.19
Farag [55]	97.23	-	-	-	-	-	-
Berrahou et al. [57]	**98.20**	98.74	99.57	65.41	69.25	97.39	90.38
HS+RR+HDF	98.12	98.47	99.68	**69.40**	59.65	97.79	89.74

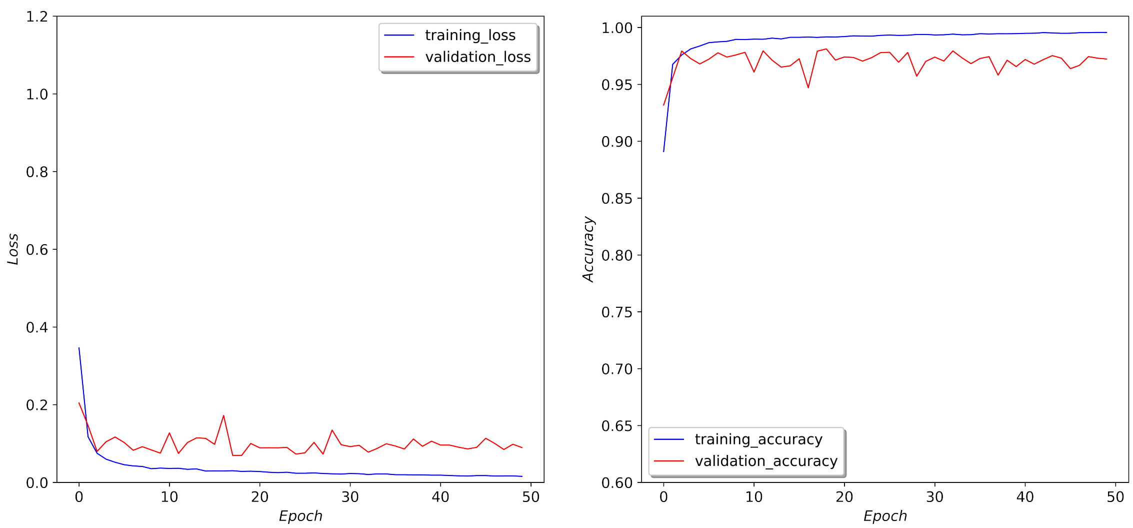

Figure 7. Accuracy and loss curves of HS+RR+HDF model using DS1 as training set and Incart dataset as the test set for the inter-patient paradigm (3 classes).

4.4. Computational Complexity Analysis

The computational complexity of our method is influenced by the individual complexities of its preprocessing, handcrafted feature extraction, automated feature extraction, and classification using the Attention-based CNN model. While many of these components demonstrate linear or near-linear complexity, the automated features extraction and classification phases, governed by the CNN model, may introduce significant complexities depending on the architecture and training process of the model. Understanding the computational complexity of each component is crucial for assessing the efficiency and scalability of our method, particularly when dealing with large datasets or real-time applications. However, as highlighted in section 3.3, it is important to note that the complexities of the other components are negligible compared to the complexity of the Attention-based CNN model.

Regarding computational complexity analysis, we have computed and presented fundamental metrics for our model configuration in Table 5, which includes the total count of layers and trainable parameters.

When assessing complexity, our model exhibits significantly lower complexity in contrast to the adversarial-based CNN model mentioned in [41]. The latter comprises seven convolutional layers and three attention modules. This simplification makes our model a favorable option for deployment on edge devices, especially when compared to the model mentioned in the referenced study.

In comparison to the model specified in [40], the total parameter count reported for this model is 1,395,648, significantly higher than the parameter count of our proposed model. Both the classification results and the model's size strongly advocate for the superiority of our proposed classifier over this referenced model. Compared to the models specified in [54, 55], these models have total parameters substantially less than our proposed model.

When comparing our proposed model with statistical classifier models such as Linear Discriminant Analysis [7], Support Vector Machine ensembles [53, 63], XGBoost classifier [31], as well as CNN-based approaches [42, 44, 45, 57, 58], hybrid CNN-LSTM deep learning models [43, 59], and attention mechanisms with Convolutional Neural Networks and LSTM [56, 60, 61], it's apparent that the computational complexity of our model is higher than traditional machine learning-based approaches but comparable to CNN-based methods.

Table 5. Computational complexity of existing models.

Model	No. of layers	No. of trainable parameters
Sellami et al. [40]	9	1,395,648
T. Wang et al. [42]	5	26.533
Zahid et al. [54]	4	23.619
Farag [55]	2	1245
Berrahou et al. [57]	5	16.693
Our proposed model	9	24.288

4.5. Model Explainability

The concept of explainability revolves around elucidating the diverse patterns within the signal that contribute to the classification of various types of arrhythmias. The aim is to offer a validation tool designed specifically for healthcare professionals. Model explainability is achieved through GradCAM, introduced in [64], which involves three steps: computing the gradient score of the final layer, determining neuron importance weights, and generating the saliency map. This map visually illustrates the significance of different spatial regions in predicting the class. In 1D Grad-CAM visualizations, ECG segments are depicted as colormaps transitioning from blue to red, indicating low to high importance. These visualizations provide insights into the relevance of specific points or areas within the ECG signal for predicting arrhythmia classes. For better model explainability, we deployed a 1-D variant of Grad-CAM++ [65], an improved iteration of Grad-CAM. Figure 8 illustrates some local Grad-CAM++ visualizations for normal (N), supraventricular ectopic (S), and ventricular ectopic (V) beat signals. The results suggest that the QRS complex exhibits the highest gradient value among normal signals, thereby emerging as the most crucial distinguishing feature for classifying normal beats. Conversely, the PQ interval and ST segment appear less significant. For supraventricular ectopic rhythms (S), attention shifts to the lower part of the RS segment. Regarding ventricular ectopic rhythms (V), the saliency maps primarily highlight the QRS complex and the latter part of the ST segment, emphasizing their critical role in classification, with importance percentages ranging from 80% to 100%.

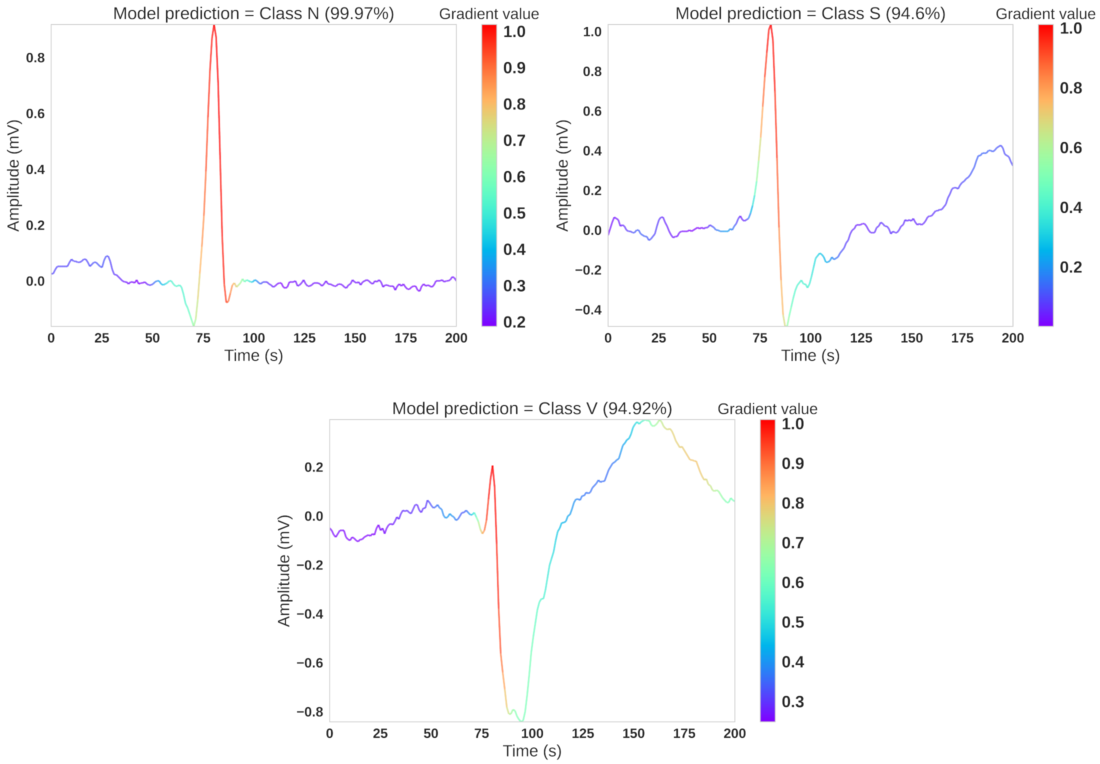

Figure 8. 1D-Grad-CAM++ visualization of ECG segment.

5. Discussion and Conclusions

In this paper, we have introduced a new method for inter-patient ECG classification encompassing four types of cardiac arrhythmias. Our approach combines the Higuchi Dimensional Fractal (HDF) and four RR intervals extracted from preprocessed ECG data, employing a 1D Convolutional Neural Network (CNN) and an attention block that encourages the model to focus on the most informative ECG signals. By computing HDF with optimal k_{max}, we utilized it as a feature set to improve accuracy. We validate the effectiveness of our proposed method using the MIT-BIH arrhythmia dataset and evaluate its generalization capability using the INCART dataset. The experimental results reveal remarkable performance, when tested on the MIT-BIH arrhythmia dataset, the method attained accuracy rates of 98.75% for three classes and 97.96% for four classes under the inter-patient paradigm. Furthermore, it exhibited precision levels of 94.58% for supraventricular ectopic beats(S) and 97.22% for ventricular ectopic beats (V) in the three-class classification. Additionally, when assessing its generalization ability on the INCART dataset, trained and validated using the MIT-BIH DS1 and DS2 datasets, our method achieved an average accuracy of 98.12% for three classes. These results highlight our method's efficacy in detecting S and V arrhythmia cases, surpassing existing methods in inter-patient analysis.

While our method demonstrates impressive performance compared to existing approaches, it encounters challenges in accurately identifying the F arrhythmia class within the inter-patient paradigm. This issue stems from the limited data representation for these minority arrhythmia types in the MIT-BIH dataset. To address this limitation and enhance the overall effectiveness of our method, we intend to leverage data augmentation techniques and integrate novel shape descriptors in a future version of our approach. Our future objectives also include analyzing large and multi-label datasets to evaluate and improve the model's generalization capabilities. We intend to increase the model's complexity and adaptability in future projects, aiming to optimize neural network performance, enhance clinical benefits, and reduce the model's time complexity through the use of various optimizers, normalization techniques, and regularization methods.

The method proposed in this paper holds considerable promise for use in medical research and clinical practice. Originally devised for arrhythmia detection, its adaptability extends to a range of areas, such as early detection of cardiac ischemia, assessment of cardiac function, prediction of cardiovascular events, and analysis of heart rate variability. Moreover, it holds promise for identifying congenital heart anomalies and sleep disorders like obstructive sleep apnea. Furthermore, there is potential to extend the model's functionality to monitor fetal heart rate. These perspectives highlight this approach's substantial potential to enhance individuals' health and overall well-being.

REFERENCES

1. Health Organization, *Cardiovascular diseases world Health Organization*, Switzerland, 2021.
2. Z. Zhang, J. Dong, X. Luo, K.S. Choi, and X. Wu, *Heartbeat classification using disease-specific feature selection.*, Comput. Biol. Med., vol.46, pp. 79–89, 2014.
3. ANSI-AAMI, *Testing and reporting performance results of cardiac rhythm and st segment measurement algorithms*, American National Standards Institute, Inc. (ANSI), Association for the Advancement of Medical Instrumentation (AAMI), ANSI/AAMI/ISO, 1998–2008.
4. P. de Chazal, M. O'Dwyer, and R. B. Reilly, *Automatic classification of heartbeats using ECG morphology and heartbeat interval features*, IEEE Trans. Biomed. Eng., vol. 51, no. 7, pp. 1196–1206, Jul. 2004.
5. R. J. Martis et al., *Computer aided diagnosis of atrial arrhythmia using dimensionality reduction methods on transform domain representation*. Biomedical Signal Processing and Control 13(1):295–305, 2014
6. R. J. Martis et al., *Application of higher order statistics for atrial arrhythmia classification*. Biomedical Signal Processing and Control 8(6):888–900, 2013
7. C.C. Lin, C.M. Yang, *Heartbeat Classification Using Normalized RR Intervals and Morphological Features*. Mathematical Problems in Engineering, pages 1-11, 2014.
8. T.Li, and M. Zhou, *ECG classification using wavelet packet entropy and random forests*. Entropy,18(8), (2016)
9. K. Luo, et al, *Patient-specific deep architectural model for ECG classification*. Journal of Healthcare Engineering, 2017. (2017)
10. Q Sun, Q Wang, B Ji, W Wu, W Huang, C Wang, *The Cardiodynamicsgram Based Early Detection of Myocardial Ischemia Using the Lempel-Ziv Complexity*. IEEE Access; 2020.8:207894–207904.
11. A. Scarciglia, V. Catrambone, C. Bonanno, G.Valenza, *Quantifying partition-based Kolmogorov-Sinai Entropy on Heart Rate Variability: a young vs. elderly study*. 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2021, pp.5469-5472.
12. BJ West, R Zhang, AW Sanders, S Miniyar, JH Zuckerman, BD Levine, *Fractal actuations in cardiac time series*, Phys A 270:552566, 1999.
13. S.S. Xu, M.W. Mak, C.C. Cheung, *Towards end-to-end ECG classification with raw signal extraction and deep neural networks*. IEEE J. Biomed. Health Inform. 2018, 23, 1574–1584.
14. U.R. Acharya, S.L. Oh, Y.Hagiwara, J.H.Tan, M. Adam, A. Gertych, R.S. Tan, *A deep convolutional neural network model to classify heartbeats*. Comput. Biol. Med. 2017, 89, 389–396.
15. E.J.S. Luz, W.R. Schwartz, G. C Amara-Chavez, D. Menotti, *ECG-based heartbeat classification for arrhythmia detection: A survey*, Comput. Methods Programs Biomed. 2016, 127, 144–164.
16. Q. Yao, R.Wang, X. Fan, J. Liu, and Y.Li, *Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time incremental convolutional neural network*. Inf. Fusion, (2020),53, 174–182.
17. J.Zhang, A. Liu, M. Gao, X.Chen, X.Zhang, and X. Chen, *ECG based multi-class arrhythmia detection using spatiotemporal attention-based convolutional recurrent neural network*. Artif. Intell. Med.(2020) 106:101856.
18. C.-H. Hsieh, Y.-S. Li, B.-J. Hwang, C.-H. Hsiao, *Detection of Atrial Fibrillation Using 1D Convolutional Neural Network*, Sensors 2020, 20,2136.
19. M. Slonim, T. Slonim, and E. Ovsyscher, *The use of simple fir filters for filtering of ECG signals and a new method for post-filter signal reconstruction*, Proceedings of Computers in Cardiology Conference, 1993.
20. V.X. Afonso, W.J. Tompkins, T.Q. Nguyen, and S. Luo, *ECG beat detection using filter banks*, IEEE Trans. Biomed. Eng., vol. 46, no. 2, pp. 192–202, Feb. 1999.
21. J. Pan, W.J. Tompkins, *A real-time QRS detection algorithm*. IEEE Trans. Biomed. Eng. 1985, BME-32, 230–236.
22. B. Chandrakar, O. Yadav, V. Chandra, *A survey of noise removal techniques for ECG signals*. Int. J. Adv. Res. Comput. Commun. Eng. 2013, 2, 1354–1357.
23. V. Gupta, M. Mittal, *QRS Complex Detection Using STFT, Chaos Analysis, and PCA in Standard and Real- Time ECG datasets.* J. Inst. Eng. Ser. B 2019, 100,489–497
24. Y. Kutlu, D. Kuntalp, *Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients*. Comput. Methods Programs Biomed. 2012, 105, 257–267.
25. M. Lagerholm, C. Peterson, G.Br accini, L. Edenbrandt, L. Sornmo, *Clustering ECG complexes using Hermite functions and self-organizing maps*. IEEE Trans. Biomed. Eng. 2000, 47, 838–848.
26. C. Ye, B.V.K.V. Kumar, and M.T. Coimbra, *classification using morphological and dynamic features of ECG signals*. IEEE Trans. Biomed. Eng. 2012, 59, 2930–2941.

27. UR. Acharya, Fujita, H. Adam,M. Oh, S.L.; Hong,TJ.; Sudarshan, VK. and Koh.J EW *Automated characterization of arrhythmias using nonlinear features from tachycardia ecg beats*. In 2016 IEEE international conference on systems, man, and cybernetics (SMC), pages 000533–000538. IEEE, 2016.
28. S. Sahoo, M. Dash, S. Behera, S. Sabut . *Machine Learning Approach to Detect Cardiac Arrhythmias in ECG Signals: A Survey*. IRBM. 41 (4), 185-194, 2020.
29. W-H. Jung, and S-G. Lee. *An arrhythmia classification method in utilizing the weighted KNN and the fitness rule*. IRBM 38.3 (2017): 138-148.
30. F. A. Elhaj, N. Salim, A. R. Harris, T.T. Swee, and T. Ahmed, *Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals*. Computer Methods and Programs in Biomedicine 127:52–63, 2016.
31. H. Shi, H.Wang, Y. Huang, L. Zhao, C. Qin, C. Liu, *A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification*. Comput. Methods Programs Biomed. 2019, 171, 1–10.
32. S. Chen,W. Hua, Z. Li , J.Li , X.Gao, *Heartbeat classification using projected and dynamic features of ECG signal*. Biomedical Signal Processing and Control 2017 31:165–173.
33. H. Dang, M. Sun, G. Zhang, X. Zhou, Q. Chang, X. Xu, *A novel deep convolutional neural network for arrhythmia classification*. In Proceedings of the 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), Shiga, Japan, 26–28 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 7–32.
34. S.L. Oh, E.Y.K. Ng, R.S. Tan, and U.R. Acharya, *Automated beat-wise arrhythmia diagnosis using modified U-net on extended electrocardiographic recordings with heterogeneous arrhythmia types*. Comput. Biol. Med. 2019, 105, 92–101.
35. E.D. Ubeyli, *Combining recurrent neural networks with eigenvector methods for classification of ECG beats*. Digit. Signal Process. 2009;19:320–329.
36. A. M. Shaker, M. Tantawi, H. A. Shedeed, and M. F. Tolba, *Generalization of convolutional neural networks for ECG classification using generative adversarial networks*, IEEE Access, vol. 8, pp. 35592–35605, 2020.
37. O. Yildirim, U.B. Baloglu, R.S. Tan, E.J. Ciaccio, U.R. Acharya, *A new approach for arrhythmia classification using deep coded features and LSTM networks*. Comput. Methods Programs Biomed. 2019, 176, 121–133.
38. O. Yildirim, *A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification*. Comput. Biol. Med. 2018, 96, 189–202.
39. S.L. Oh, E.Y.K. Ng, R.S. Tan, and U.R. Acharya, *Automated diagnosis of arrhythmia using a combination of CNN and LSTM techniques with variable length heartbeats*. Comput. Biol. Med. 2018, 102, 278–287.
40. A. Sellami and H. Hwang, *A robust deep convolutional neural network with batch-weighted loss for heartbeat classification*, Expert Syst. Appl., vol. 122, pp. 75–84, May 2019.
41. J. Zhang, A. Liu, D. Liang, X. Chen, M. Gao, *Interpatient ECG Heartbeat Classification with an Adversarial Convolutional Neural Network*. J. Healthc. Eng. 2021.
42. T.Wang, C. Lu, Y. Sun, M. Yang, C. Liu, C. Ou, *Automatic ECG classification using continuous wavelet transform and convolutional neural network*. Entropy 2021.
43. H. Shi, C. Qin, D. Xiao, L. Zhao, C. Liu, *Automated heartbeat classification based on deep neural network with multiple input layers*, Knowl.-Based Syst. 188 (2020) 105036.
44. J. Niu, Y.Tang, Z.Sun, W.Zhang. *Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks*. IEEE J Biomed Health Inform 2020;24(5):1321–32.
45. F. Li, Y. Xu, Z.Chen, Z.Liu. *Automated heartbeat classification using 3-d inputs based on convolutional neural network with multi-fields of view*. IEEE Access 2019;7:76295–304.
46. L. Guo, G.Sim , B.Matuszewski. *Inter-patient ECG classification with convolutional and recurrent neural networks*. Biocybern Biomed Eng 2019;39(3):868–79.
47. 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: Components of a new research resource for complex physiologic signals*, Circulation, vol. 101, no. 23, pp. e215–e220, Jun. 2000.
48. 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/Jun. 2001
49. J. He, J. Rong, L. Sun, H. Wang, Y. Zhang, J. Ma, *A framework for cardiac arrhythmia detection from IoT-based ECGs*. World Wide Web 2020, 23, 2835–2850.
50. T. Wang, C. H. Lu, M. Yang, F.Hong, and C. Liu, *A hybrid method for heartbeat classification via convolutional neural networks, multilayer perceptrons, and focal loss*, Peerj Computer Science, vol. 6, Article ID e324, 2020.
51. T. Higuchi, *Approach to an irregular time series on the basis of the fractal theory*, Phys. D Nonlinear Phenom.vol. 31, no. 2, pp. 277–83, 1988.
52. M. Ilse, J.M. Tomczak, M. Welling, *Attention-based deep multiple instance learning*, CoRR abs /1802.04712 (2018)
53. G. Garcia, G. Moreira, D. Menotti, E. Luz, *Inter-patient ECG heartbeat classification with temporal VCG optimized by PSO*. Sci.Rep. 2017, 7, 1–11.
54. M.U. Zahid, S. Kiranyaz, M. Gabbouj, *Global ECG classification by self-operational neural networks with feature injection*. IEEE Trans. Biomed. Eng., 1–12,2022.
55. M.M. Farag, *A Tiny Matched Filter-Based CNN for Inter-Patient ECG Classification and Arrhythmia Detection at the Edge*. Sensors 2023, 23, 1365.
56. P. Xu, H. Liu, X. Xie, S. Zhou, M. Shu, Y. Wang, *Interpatient ECG arrhythmia detection by residual attention CNN,* Comput. Math. Methods Med. 2022 (2022),
57. N. Berrahou, A. El Alami, A. Mesbah, R. El Alami, and A.Berrahou, *Arrhythmia detection in inter-patient ECG signals using entropy rate features and RR intervals with CNN architecture.* Computer Methods in Biomechanics and Biomedical Engineering, 1–20.(2024)
58. G. Wang, M. Chen, Z. Ding, J. Li, H. Yang, P. Zhang, *Inter-patient ECG arrhythmia heartbeat classification based on unsupervised domain adaptation*. Neurocomputing 2021, 454, 339–349.

59. Y. Xia, Y. Xu, P. Chen, J. Zhang, Y. Zhang, *Generative adversarial network with transformer generator for boosting ECG classification*, Biomed. Signal Process. Control 80 (2023) 104276.
60. F. Zhou, Y. Sun, Y. Wang, *Inter-patient ECG arrhythmia heartbeat classification network based on multiscale convolution and FCBA*, Biomedical Signal Processing and Control, 90,2024,105789.
61. A. Chen, et al., *Multi-information fusion neural networks for arrhythmia automatic detection,* Comput. Methods Programs Biomed. 193 (2020) 105479.
62. M Llamedo, JP Martinez, *Heartbeat classification using feature selection driven by database generalization criteria*. IEEE Trans Biomed Eng. (2011) 58:616–625.
63. J. He, J. Rong, L. Sun, H. Wang, Y. Zhang, *An Advanced Two-Step DNN-Based Framework for Arrhythmia Detection*. Lect. Notes Comput. Sci. 2020, 12085, 422–434.
64. Selvaraju, M Cogswell, A Das, R Vedantam, D Parikh, D Batra. *Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization*. In: 2017 IEEE International Conference on Computer Vision (ICCV);2017. p. 618–626.
65. A. Chattopadhay et al., *-cam++: Generalized gradient-based visual explanations for deep convolutional networks*, in: 2018 IEEE winter conference on applications of computer vision (WACV), IEEE, 2018.