

A Novel Inter Patient ECG Arrhythmia Classification Approach with Deep Feature Extraction and 1D Convolutional Neural Network

Mohamed Elmehdi Ait Bourkha^{1*}, Anas Hatim², Dounia Nasir³, Said El Beid⁴, Assia Sayed Tahiri⁵

Information Technology and Modeling Team (TIM), National School of Applied Sciences (ENSA) of Marrakech, Cadi Ayyad University (UCA), Marrakech, Morocco^{1, 2, 3, 5}

Control and Computing for Smart Systems and Green Energy (CISIEV), ENSA of Marrakech, UCA, Marrakech, Morocco⁴

Abstract—The World Health Organization (WHO) sheds light on the escalating prevalence of heart diseases, foreseeing a substantial rise in the years ahead, impacting a vast global population. Swift and accurate early detection becomes pivotal in managing severe complications, underscoring the urgency of timely identification. While Ventricular Ectopic Beats (V) might initially be considered normal, their frequent occurrence could serve as a potential red flag for progressing to severe conditions like atrial fibrillation, Ventricular Tachycardia, and even cardiac arrest. This accentuates the need for developing an automated approach for early detection of cardiovascular diseases (CVD). This paper presents a novel method to classify arrhythmias. Leveraging the Wavelet Scattering Transform (WST) to extract morphological features from Electrocardiogram heartbeats (ECG), these features seamlessly integrate into a 1D Convolutional Neural Network (CNN). The CNN is finely tuned to distinguish between V, Supraventricular Ectopic Beats (S), and Non-Ectopic Beats (N). Our model's performance surpasses state-of-the-art models, boasting precision, sensitivity, and specificity of 94.56%, 97.26%, and 99.54% for V, and 99.25%, 98.65%, and 93.26% for N. Remarkably, it achieves 68.01% precision, 77.75% sensitivity, and 99.14% specificity for S.

Keywords—*Electrocardiogram (ECG); Cardiovascular Diseases (CVD); Wavelet Scattering Transform (WST); Convolutional Neural Network (CNN)*

I. INTRODUCTION

According to the World Health Organization (WHO), cardiovascular diseases (CVD) stand as the foremost cause of global mortality, contributing to approximately 17.9 million deaths annually. This staggering toll is predominantly attributed to coronary heart disease, encompassing heart attacks, and Cerebrovascular illnesses, including strokes, which collectively account for 80% of these fatalities. The demographic most vulnerable to these health challenges comprises individuals in their middle and elderly years [1]. Clinical research in the preceding century has significantly advanced to facilitate the early detection of CVD. During the historical period when the electrical nature of the heart was widely acknowledged, a notable challenge existed due to the absence of tools for its systematic study. It is noteworthy that the exploration of electricity in the medical domain had commenced nearly two centuries before Einthoven's pivotal contributions. Pioneers such as: Gilbert with his work "De Magnete" in 1600, Bacon through "Novum Organum" in 1620,

and Browne who coined the term electricity in the mid-seventeenth century within "Pseudodoxia Epidemica" in 1646, had laid the foundation for understanding electrical phenomena in the context of medicine.

Despite this early groundwork, it was only with Einthoven's groundbreaking efforts that a tangible method, namely the EKG or ECG, was developed to specifically delve into the intricate electrical dynamics of the heart [2]. In the year 1893, Einthoven not only introduced the term "electrocardiogram" but also demonstrated notable advancements by enhancing the electrometer. His significant contributions included the introduction of a correction formula, a pivotal innovation that enabled the differentiation of five distinct deflections [3-4]. These deflections, denoted by the names PQRST, were assigned based on the Cartesian nomenclature.

Comprising P, QRS, and T waves, along with an array of intervals, the electrocardiogram (ECG) serves as a comprehensive diagnostic tool. This intricate waveform not only facilitates the analysis of fundamental parameters such as heart rate but also provides invaluable insights into the intricate conditions and potential risks associated with the heart's functionality [5].

By examining these distinctive components and intervals, clinicians can glean a nuanced understanding of cardiac health, enabling them to address and mitigate potential issues proactively. Determining the signal's classification into the normal range or otherwise relies crucially on assessing the amplitude measured in millivolts and the intervals expressed in milliseconds [6].

Sinus rhythm, denoting the typical heart rhythm, is characterized by the orchestrated propagation of triggering impulses originating from the sinoatrial node. This synchronized transmission extends seamlessly throughout the four chambers of the heart, ensuring a harmonious and coordinated cardiac activity [7]. The underlying cause of Cardiac arrhythmia (ARR) stems from the disruption of proper electrical impulses that orchestrate the coordinated beats of the heart (too slow or too fast) [7]. It is noteworthy that certain ARRs, if left unaddressed, pose a serious threat, potentially leading to sudden cardiac death. Hence, understanding and addressing these dangerous arrhythmic conditions becomes paramount in safeguarding cardiovascular health.

ARRs encompassing four prevalent types: extra beats, supraventricular tachycardia, ventricular ARR, and bradyarrhythmias [8-9]. Premature (extra) beats include both premature atrial contractions (PAC) and premature ventricular contractions (PVC). Supraventricular ARR initiates in the atria, characterized by elevated heart rates, they include atrial fibrillation (AF), atrial flutter, paroxysmal supraventricular tachycardia (PSVT), and Wolff-Parkinson-White syndrome. Within the realm of ventricular ARR lie Ventricular flutter, Ventricular tachycardia (VT), and Ventricular fibrillation (VF).

These potentially life-threatening conditions necessitate immediate intervention, often requiring the prompt administration of a defibrillator shock to safeguard and preserve life. Heart block, a condition that can manifest in the atrioventricular node or the HIS Purkinje system, it leads to an irregular slowed heartbeat, potentially necessitating the use of a pacemaker for treatment. The condition is observed on either the left or right side of the ventricles, identified as right bundle branch block (RBBB) or left bundle branch block (LBBB).

Precise interpretation of ECG signals holds the potential to avert the progression of chronic heart conditions to irreversible stages. However, the manual interpretation and analysis of ECG pose a formidable challenge [10-11]. Moreover, beyond the challenge of manual interpretation, the analysis of long-term ECG signals is a time-intensive endeavor, susceptible to human errors that may compromise the accuracy of assessments [12-13-14, 15].

In recent decades, researchers have responded to the challenges by developing automated models aimed at the detection and in-depth analysis of long-term ECG signals. This innovative approach not only addresses the limitations associated with manual interpretation but also heralds a paradigm shift in enhancing the efficiency and accuracy of cardiac signal assessment over extended periods. These techniques solely rely on advancements in machine learning and deep learning [16-17, 18], which have seen significant progress in recent years, thanks to the continuous development of computer systems and technology.

In the ensuing sections, relevant literature was explored, outlining automated heartbeat classification models. Subsequently, in the methodology, the meticulous preparation and processing of our data were unfolded, the method employed for feature extraction was elucidated, and a comprehensive explanation of the classifier utilized was provided; including its parameters, along with the evaluation metrics employed to assess the efficacy of our proposed model.

Moving forward, the results and discussion section expounded on our findings, drawing comparisons with state-of-the-art systems. Lastly, the conclusion section summarized our results, highlighted the limitations inherent in this study, and proposed avenues for future research to address these constraints.

II. RELATED WORKS

A. Methodology Overview in Previous Studies

This section introduces state-of-art automated inter-patient ECG heartbeat classification models, adhering to the

Association for the Advancement of Medical Instrumentation (AAMI) recommendation [19]. These models categorize ECG heartbeats into five types according to the AAMI standard: Normal (N), Supraventricular ectopic beats (S), Ventricular ectopic beats (V), Fusion beats (F), and Unknown beats (Q), as outlined in Table I.

TABLE I. AAMI HEARTBEAT CLASSES

AAMI Classes	MIT-BIH heartbeats
N	Normal (N)
	Right bundle branch block beats (RBBB)
	Nodal (junctional) escape beats (j)
	Left bundle branch block beats (LBBB)
	Atrial escape beats (e)
S	Aberrated atrial premature beats (a)
	Atrial premature contraction (A)
	Supraventricular premature beats (S)
	Nodal (junctional) premature beats (J)
V	Premature Ventricular contraction (PVC)
	Ventricular escape beats (E)
	Flutter wave (!)
F	Fusion of ventricular and normal beat (F)
Q	Paced beat (/)
	Fusion of paced and normal beat (f)
	Unclassified beat (Q)

Sellami et al. [20] introduced a deep CNN, leveraging state-of-art deep learning techniques for precise heartbeat classification. They advocated for a batch-weighted loss function to effectively address class imbalances, wherein the loss weights dynamically adjust based on the changing class distribution in each batch. Additionally, they proposed the utilization of multiple heartbeats to enhance the classification of the five heartbeat classes. The evaluation of their proposed approach on inter-patient data from the MIT-BIH arrhythmia database yielded results, their model achieving an accuracy, precision, sensitivity, and specificity of 88.34%, 45.25%, 90.90%, and 88.51%, respectively.

Li et al. [21] employed the overlapping segmentation method to divide ECG signals from the MIT-BIH database into 5-second segments, addressing class imbalances by re-labeling these segments. Subsequently, discrete wavelet transform (DWT) was applied for denoising, and a deep residual CNN was employed for ARR classification. They incorporated the focal loss function. Their proposed method demonstrated a sensitivity, precision, and specificity of 94.54%, 93.33%, and 80.80% for class N. For class S, the model achieved 35.22% sensitivity, 65.88% precision, and 98.83% specificity, while for class V, the method exhibited 88.35% sensitivity, 79.86% precision, and 94.92% specificity.

Garcia et al. [22] presented an innovative approach wherein they proposed an ECG representation based on vectorcardiogram, termed temporal vectorcardiogram. They employed a complex network for feature extraction and fine-

tuned an SVM classifier using the particle swarm optimization algorithm. The results of their approach, applied to inter-patient analysis on the MIT-BIH arrhythmia database, demonstrated a precision of 53% for class S and 87.3% for class V.

Wang et al. [23] introduced an automated ECG classification method that relies on Continuous Wavelet Transform (CWT) and CNN. The CWT is employed to decompose ECG signals, yielding distinct time-frequency components. Subsequently, the CNN is utilized to extract features from the 2D-scalogram composed of these time frequency components. 4 RR interval features are extracted and combined with the CNN features and inputted into a fully connected layer. The proposed model demonstrated a sensitivity of 99.04%, precision of 98.04%, and specificity of 87.95% for class N. For class S, the model achieved a sensitivity of 70.75%, precision of 77%, and specificity of 99.51%. Additionally, for class N, the sensitivity, precision, and specificity were 94.35%, 95.32%, 99.45%, respectively.

Takalo et al. [24] introduced a method for inter-patient ECG heartbeat classification, leveraging a deep CNN. Their proposed approach demonstrated commendable performance, achieving a sensitivity of 92% and precision of 97% for class N. Additionally, for class S, the method attained a sensitivity of 62% and precision of 56%. Finally, in the case of class V, their approach yielded a sensitivity of 89% and a precision of 51%.

Junaid et al. [25] proposed a 1D self-organized operational neural network for the inter-patient classification of ECG heartbeats from the MIT-BIH arrhythmia database. Their method exhibited an overall accuracy of 95.99% across: N, S, and V. Specifically, for class N, the proposed method achieved a sensitivity of 98.48%, precision of 97.39%, and specificity of 76.82%. In the case of class S, they attained a sensitivity of 44.01%, precision of 64.50%, and a specificity of 99.01%. Lastly, for class V, the method demonstrated a sensitivity of 92.96%, precision of 89.62%, and a specificity of 99.22%.

He et al. [26] introduced a new method, multi-level unsupervised domain adaptation framework (MLUDAF), for diagnosing ARRs. They used the spatial pyramid pooling residual (ASPP-R) module to extract spatio-temporal features and employed the graph convolutional network (GCN) module for data structure features. In domain adaptation, they aligned domains, semantics, and structures. Testing on MIT-BIH yielded a 96.8% overall accuracy. Notably, the model achieved 97.8% sensitivity and 99.5% precision for N class, 89.2% sensitivity and 90.4% precision for V class, and 90.2% sensitivity and 53.2% precision for S class.

It's crucial to highlight those previous studies have largely overlooked classes F and Q due to their minimal representation (less than 1%) in the MIT-BIH dataset. As these classes make an insignificant contribution to overall performance, they were excluded. Including them in model training would lead to an imbalanced dataset, adversely impacting classification results. Thus, this research focuses exclusively on classes N, S, and V for a more effective analysis.

B. Limitations and Gaps in Previous Studies

It's noteworthy that Wang et al. [23] achieved the best results in distinguishing N, S, and V classes, with an overall

accuracy of 97.68%. However, their approach fell short, particularly in classifying S with a sensitivity below 70.75% and limited precision at 77%. Similar limitations were observed in class V, where sensitivity did not exceed 94.35%, and precision reached only 95.32%. This underscores the need for a novel approach to enhance predictions for N, S, and V.

In previous studies, various methods and approaches based on intelligent models have been employed to classify N, S, and V classes, yet the classification results remain inadequate for clinical applications. Therefore, in this research, we aim to investigate the efficacy of a deep feature extraction method using WST, complemented by a second stage feature extraction through 1D CNN. Our objective is to improve prediction and classification results for the three classes.

The rationale for selecting the WST lies in its robust capability to extract both time-domain and frequency-domain features crucial for effectively distinguishing between the three classes. Furthermore, the choice of the 1D CNN in this study is justified by its inherent feature extraction stage, which facilitates the extraction of deeper features. Additionally, the 1D CNN has demonstrated remarkable performance in various classification tasks, underscoring its widespread applicability and effectiveness in classification endeavors.

The combination of these two techniques aims to address the limitations encountered in previous works and improve the prediction accuracy of the 3 ECG classes. This enhancement holds significant promise for advancing the field of automated detection of ARR in ECG. By integrating the WST and 1D CNN, our study seeks to surpass the constraints of prior methodologies and achieve more precise classification results.

III. MATERIALS AND METHODS

A. Database Description

The research relies on publicly available data from PhysioNet, accessible at the following link: <https://www.physionet.org/content/mitdb/1.0.0/>. The ECG readings were extracted from the MIT-BIH Arrhythmia database, comprising 48 half-hour excerpts of two channel ambulatory ECG recordings [27].

The recordings were digitized at a rate of 360 samples per second per channel, with an 11-bit resolution over a 10 mV range. The records used in this study are taken from the lead II, because lead II is positioned along the axis of the heart, making it particularly sensitive to changes in heart's rhythm that occur in this plane. So, it's an important tool for detecting ARRs.

In this study, recordings with paced beats were excluded from the MIT-BIH dataset, specifically four recordings (102, 104, 107, and 217) out of the total 48. The reason for excluding those patients from the analysis was that they had cardiac pacemakers, which had the potential to cause interference.

B. Data Preprocessing

Our study focuses on long-term ECG signals from the MIT-BIH database, aiming to classify heartbeats into three categories: N, S, and V. To achieve this, a QRS detection algorithm is essential for extracting distinct ECG heartbeats from a single recording. Numerous approaches have been

explored in the realm of QRS detection, ranging from the foundational Pan-Tompkins algorithm [28] to wavelet-based techniques [29-30], and even machine learning models [31].

In our investigation, the Pan-Tompkins algorithm was opted due to its proven accuracy in QRS detection, achieving an impressive 99.3% accuracy when tested on the MIT-BIH arrhythmia database. Additionally, the algorithm exhibits a lower computational cost compared to alternative QRS detection models. This choice aligns with our goal of efficiently and accurately discerning QRS complexes in ECG signals for subsequent heartbeat classification.

Following the QRS complex detection using the Pan-Tompkins algorithm, 70 samples before and 70 samples after each R peak were extracted, resulting in a total of 141 samples per heartbeat. Fig. 1, Fig. 2, and Fig. 3 depict three distinct types of ECG heartbeats: N, S, and V. To accommodate variations in heart rate, the difference between the post-RR interval and pre-RR interval was incorporated into the samples.

To ensure a fair comparison with previous automated heartbeat classification models, a data split strategy consistent with state-of-the-art approaches was adopted. Similar to these studies, 44 MIT-BIH recordings (excluding four with paced beats) were partitioned into two subsets. Half of the recordings (DS1) were utilized for training, while the remaining constituted the test set (DS2). The distribution of the data used in this paper is outlined in Table II. This inter-patient data split methodology enhances the development of well-generalized model, capable of effectively classifying ECG from new, unseen patients.

In this paper, the DS1 subset was employed for model training and DS2 for testing. A notable observation within DS1 revealed a substantial class imbalance, where the N class boasted 46,596 instances, while the minority classes, S and V, were represented by 1,669 and 3,799 instances, respectively. Recognizing that training on such an imbalanced dataset could skew results towards the majority class and amplify bias, we implemented the Synthetic Minority Over Sampling Technique (SMOTE) [32]. SMOTE proved instrumental in mitigating this imbalance by generating synthetic instances for the minority class. This strategic approach aimed to rectify the skewed class distribution, and foster more equitable representation of each heartbeat type during training. By addressing this imbalance, we sought to enhance the classification performance of our model, ensuring its effectiveness across all classes.

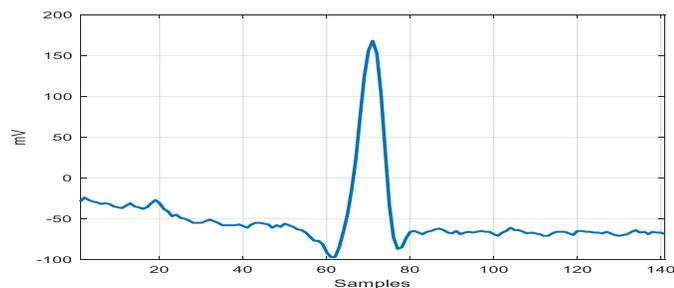


Fig. 1. Non-Ectopic beat.

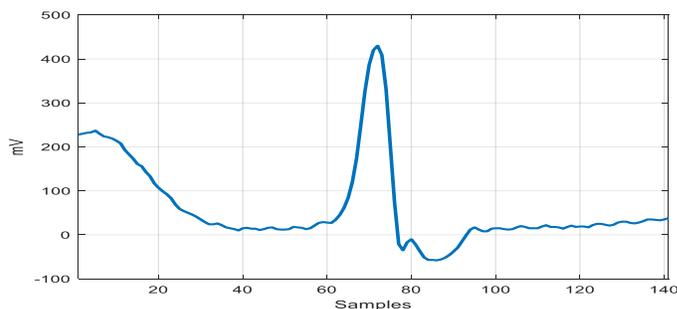


Fig. 2. Supraventricular ectopic beat.

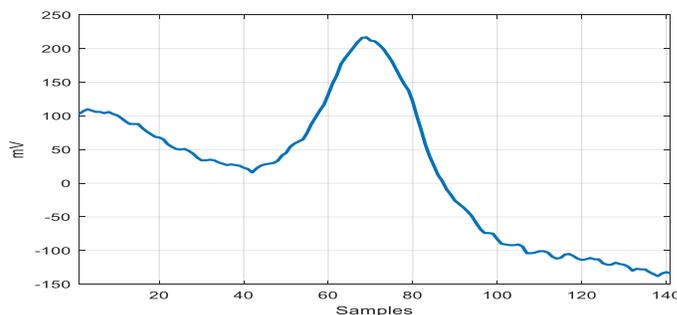


Fig. 3. Ventricular ectopic beat.

TABLE II. RECORDS AND HEARTBEAT CLASSES PARTITION

Partition	Patients ID	N	S	V
DS1	101, 106, 108, 109, 112, 114, 115, 116, 118, 119,122,124, 201, 203, 205, 207, 208,209, 215,220, 223, 230	46596	1669	3799
DS2	100,103,105, 111, 113,117, 121, 123, 200, 202,210, 212, 213,214, 214,219,221,222,228,231,232, 233,234	43478	1110	3681
Total		90074	2779	7480

C. Deep Features Extraction

To achieve accurate ECG heartbeat classification, it is crucial to employ a feature extraction technique. Two main types of features play a key role in this process: time-domain features, which capture and analyze signal amplitudes over time, and frequency-domain features, providing insights into different signal frequencies. However, traditional approaches like Fourier transform, commonly used for extracting frequency-domain features, lack temporal information, only revealing the frequencies present in the signal. Addressing this limitation, the wavelet scattering transform (WST) emerges as a solution. Fig. 4 illustrates the composition of a wavelet scattering network, featuring multiple layers where WST, akin to a CNN with cascading wavelets, is performed at each layer.

The WST offers features that exhibit stability against deformation, as well as invariance to translation and rotation. Notably, it serves as a potent technique for signal denoising and dimensionality reduction. Applied extensively in audio, image, and 1D signal analysis [33].

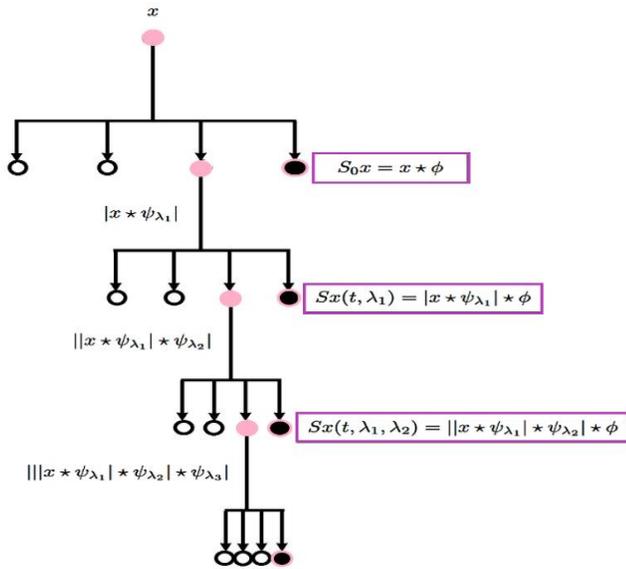


Fig. 4. Scattering network architecture

This paper introduces the Morlet analytic wavelet, offering several advantages. Firstly, being a complex wavelet, it allows the utilization of both its real and imaginary components in convolution processes. Additionally, the wavelet features a low-pass filter represented by the modulus of this wavelet, facilitating down-sampling during the scattering transform.

Moreover, Fig. 5 highlights the remarkable similarity between the shape of this complex wavelet and the QRS complex in ECG heartbeats. This similarity enhances its sensitivity in analyzing and detecting ARR within ECG signals. The expression below represents the mathematical representation of the complex wavelet employed.

$$\psi(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{t^2}{2\sigma^2}} e^{i\omega t} \quad (1)$$

In the given expression, t corresponds to time, and σ stands for the standard deviation of the Gaussian function. ω is defined as $2\pi f$, where f is the center frequency of ψ , and i is the imaginary unit. The envelope of the complex wavelet is characterized as a low-pass filter denoted as Φ in Eq. (2).

$$\Phi(t) = |\psi(t)| \quad (2)$$

The WST initiates by convolving $x(t)$, which represents the signal being analyzed, with Φ .

$$S_0x(t) = x(t) * \Phi \quad (3)$$

The wavelet function ψ and the low-pass filter Φ are specifically designed to span the entire frequency range of the signal $x(t)$. The low-pass filter introduces a form of averaging, ensuring locally invariant translation features of x . However, the initial order of the Wavelet Scattering Network (WSN) results in the loss of high frequencies, which can be restored by progressing to the subsequent order of the WSN.

In the 1st order of the WSN, an additional convolution was applied using the wavelet ψ_{λ_1} with the scale λ_1 . Here, $\psi_{\lambda_1} \in$

$\{\psi_{\lambda_k}\}_{\lambda_k \in \Delta_k}$, where, ψ_{λ_k} represents the multi-scale high-pass filter bank, and Δ_k represents the family of wavelet indices with an octave frequency resolution Q_k .

$$|W_1|x = \{S_0x(t), |x * \psi_{\lambda_1}|\}_{\lambda_1 \in \Delta_1} \quad (4)$$

The 1st order scattering coefficients result from the convolution with the low-pass filter Φ .

$$S_1x(t) = \{|x * \psi_{\lambda_1}| * \Phi\}_{\lambda_1 \in \Delta_1} \quad (5)$$

To retrieve the information (high frequencies) lost during the application of the low-pass filter, the second order of the WSN was proceed.

$$|W_2|x = \{S_1x(t), |x * \psi_{\lambda_1}| * \psi_{\lambda_2}\}_{\lambda_2 \in \Delta_2} \quad (6)$$

Obtaining the 2nd order coefficients of the WSN involves applying convolution with the low-pass filter Φ .

$$S_2x(t) = \{||x * \psi_{\lambda_1}| * \psi_{\lambda_2}| * \Phi\}_{\lambda_2 \in \Delta_2} \quad (7)$$

Similarly, we iterate through this process to determine the coefficients of the n^{th} order in the scattering network.

$$S_nx(t) = \{|\dots ||x * \psi_{\lambda_1}| * \psi_{\lambda_2}| * \dots \psi_{\lambda_n}| * \Phi\}_{\{\lambda_1, \dots, \lambda_n\} \in \{\Delta_1, \dots, \Delta_n\}} \quad (8)$$

The final scattering matrix coefficients represented as:

$$Sx(t) = \{S_0, S_1, \dots, S_n\} \quad (9)$$

In this study, only a 2nd order WSN was used. This is because 99% of the energy of the signal under analysis is preserved in this layer. Furthermore, the addition of more layers to the network would result in a loss of information about the signal and an increase in computational cost.

For the representation of a single ECG heartbeat, applying the WST yields a tensor of size 34x5, which serves as the feature matrix. The scattering coefficients undergo critical down sampling based on the bandwidth of the low-pass filters, generating five-time windows for each of the 34 scattering paths. Each row and column in the tensor correspond to a specific scattering path and time window. Furthermore, the difference between post-RR and pre-RR interval was introduced through feature fusion, oversampling this value by a factor of 34. So, the final representation of the feature matrix is size 34x6. With a total of 48,269 testing instances, this results in a tensor size of 48,269x34x6. Same process is applied to training instances.

Utilizing the WSN with an invariance scale of 0.3 seconds and employing $Q_1=8$ and $Q_2=1$ as the quality factors for the 2 filter banks. The frequency bands of the 1st and 2nd filter banks are illustrated in Fig. 6.

Fig. 7 displays the scattering coefficients obtained through the application of WST. These coefficients serve as a visual representation of the hierarchical feature extraction process. The scattering coefficients indicate the presence and distribution of significant features within the signal, offering insights into the signal's composition at various levels.

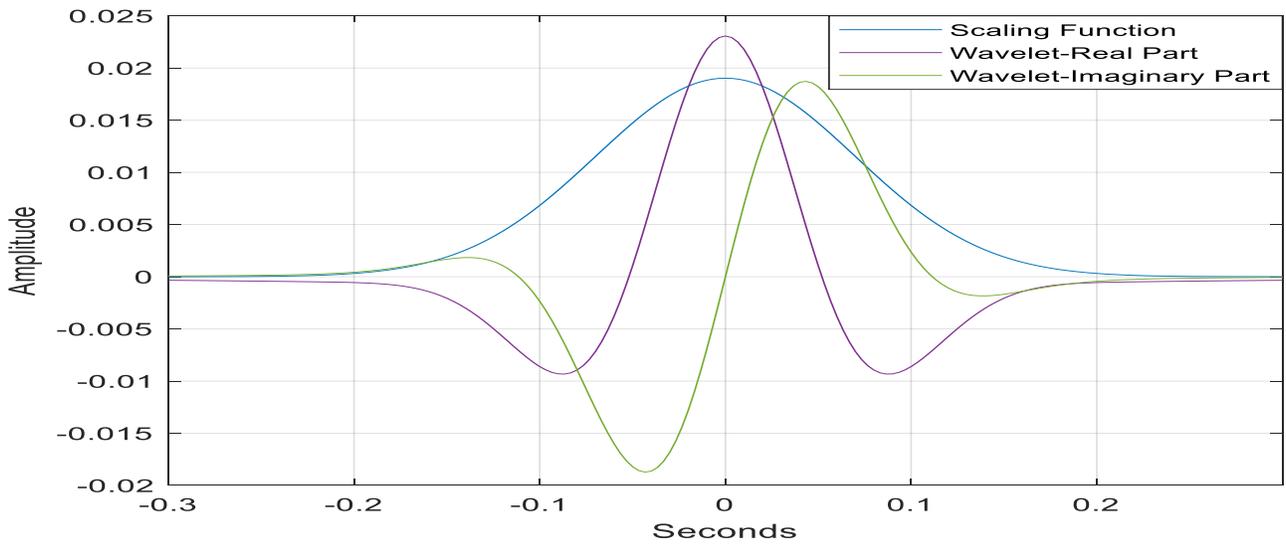


Fig. 5. Morlet Analytics wavelets.

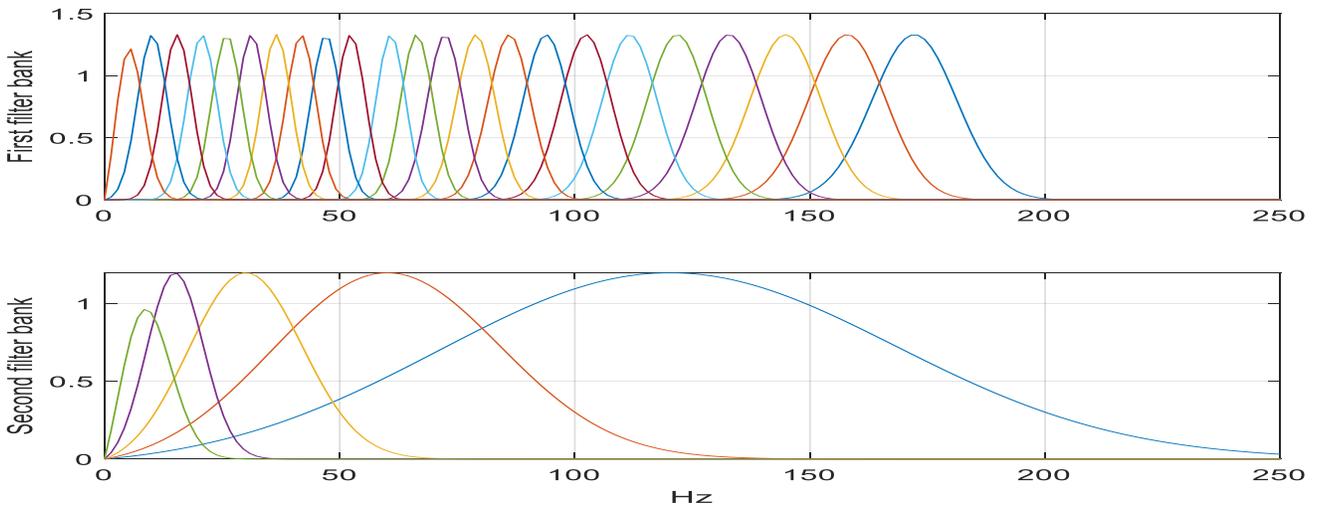


Fig. 6. Frequency bands of first and second filter banks.

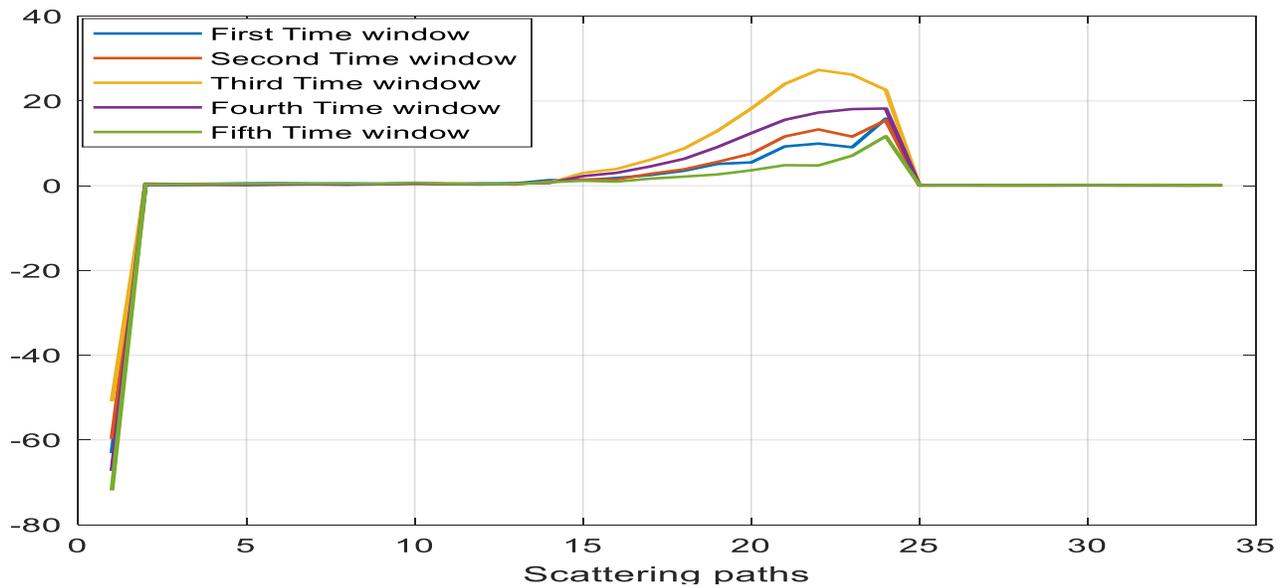


Fig. 7. Scattering coefficients for one ECG heartbeat.

At a glance, one can observe the dominant patterns and structures captured by the different orders of scattering. Higher order coefficients reveal more intricate details and relationships within the signal, while lower order coefficients highlight foundational features, analyzing the spatial distribution and patterns in these coefficients aids in interpreting the discriminative aspects and inherent characteristics of the signal under consideration. This visual representation enhances the understanding of how the WST effectively captures and organizes relevant information for further analysis.

The WSN was implemented in the MATLAB environment, with the WSN parameters configured for an invariance scale of 0.3 seconds. The sampling frequency was set at 360 Hz.

D. Classification Model

In our research endeavor, a 1D Convolutional Neural Network (1D CNN) was meticulously utilized to address the intricate task of classifying ECG heartbeats into three discernible categories. This deliberate choice of utilizing a 1D CNN stems from its widely acknowledged effectiveness in both the realms of classification and feature extraction, a critical aspect of signal analysis.

The architecture of the 1D CNN unfolds in a systematic manner, featuring two fundamental components. In the initial phase, convolution is employed with fixed-size filters to extract salient features from the signal under scrutiny. This process plays a pivotal role in enhancing the network's ability to discern subtle patterns inherent in ECG data. The subsequent infusion of non-linearity is achieved through the application of the (ReLU) activation function, a prevalent choice in this context due to its efficacy in preventing vanishing gradient.

Moving forward, the convolutional stage is followed by pooling, strategically implemented to reduce the dimensionality of the extracted features. This reduction not only aids in computational efficiency but also enhances the network's ability to focus on the most relevant information for subsequent classification steps.

The final stages of the 1D CNN involve a fully connected layer, which serves as a critical bridge between the extracted features and the subsequent classification process. The output layer, equipped with an activation function tailored for signal classification, culminates in the conclusive step of categorizing ECG signals into their respective classes.

One notable distinction in our approach is the deliberate choice of a 1D CNN over a 2D CNN. This decision is rooted in a keen consideration of computational costs, with the 1D CNN proving to be a more resource-efficient solution. This pragmatic approach underscores our commitment to not only achieving high classification accuracy but also optimizing the computational demands, making our methodology a judicious choice for real-world applications.

Our designed 1D CNN boasts a sophisticated architecture comprising a total of 11 layers. The inaugural layer, known as the sequence input layer, plays a crucial role in organizing input rows as sequences, setting the stage for subsequent treatment. Following this, our 2nd layer takes the form of a convolution layer with a filter size of 3 and 64 filters. Notably,

a fixed padding size of [2,0] was employed to enhance the convolutional process. In the 3rd layer, the non-linearity was infused into the model through the ReLU activation function. Finally, the 4th layer introduces a normalization step, a pivotal measure prevented the exploding gradient phenomenon. In the 5th layer, another convolution layer was employed with a filter size of 3 and an increased 128 filters. Accompanied by a padding size of [2,0], this convolutional stage is complemented by ReLU activation and normalization.

The convolution, ReLU, and normalization processes are represented by the following Eq. (10), Eq. (11) and Eq. (12).

$$y_i = b + \sum_{j=0}^{n-1} W_j X_{i+j} \quad (10)$$

In the output feature map, y_i represents the output at position i where b is the bias term, w_j is the weight at position j in the filter kernel, x_{i+j} refers to the input at position $i+j$ in the input feature map, and n denotes the width of the filter kernel. The ReLU function sets negatives to zero, leaving positives unaffected.

$$y_i = \max(0, x_i) \quad (11)$$

The normalization process can be modeled as follows.

$$Z = \frac{x-\mu}{\sigma} \quad (12)$$

where, μ designed the mean value and σ is the standard deviation.

In the 8th layer, global average 1D pooling is applied to decrease dimensionality. Subsequently, a fully connected layer with an output layer featuring a SoftMax activation function follows. Eq. (13) and Eq. (14) below illustrate the fully connected layer and the SoftMax activation function.

$$Z_j = \sum_{i=1}^n W_{ij} X_i + b_j \quad (13)$$

$$Y_c = \frac{e^{Z_c}}{\sum_{k=1}^3 e^{Z_k}} \quad (14)$$

where, Y_c is the predicted probability for class c after applying the SoftMax activation function.

Fig. 8 illustrates the architecture of our 1D CNN designed for the classification process, while Table III displays the specific parameters set for training the model.

E. Evaluation Metrics

In assessing the efficacy of our model and conducting a comprehensive comparison with state-of-art models, it becomes imperative to employ a spectrum of performance metrics. Key measures, including accuracy, sensitivity, specificity, precision, and the F1 score, play pivotal roles in providing a nuanced evaluation of the model's capabilities.

$$\text{Accuracy (ACC)} = \frac{TP+TN}{TP+TN+FP+FN} \quad (15)$$

$$\text{Precision (PPV)} = \frac{TP}{TP+FP} \quad (16)$$

$$\text{Recall or Sensitivity (SEN)} = \frac{TP}{TP+FN} \quad (17)$$

$$\text{Specificity (SPE)} = \frac{TN}{TN+FP} \quad (18)$$

$$F1 \text{ score (F1)} = \frac{2 * \text{precision} * \text{sensitivity}}{\text{precision} + \text{sensitivity}} \quad (19)$$

F. System Description

All algorithms were implemented using MATLAB version R-2021b on a Windows server. The system used for execution had an Intel Core i5 6300U processor running at 2.40 GHz, equipped with 12 GB of RAM, and operated on a 64-bit.

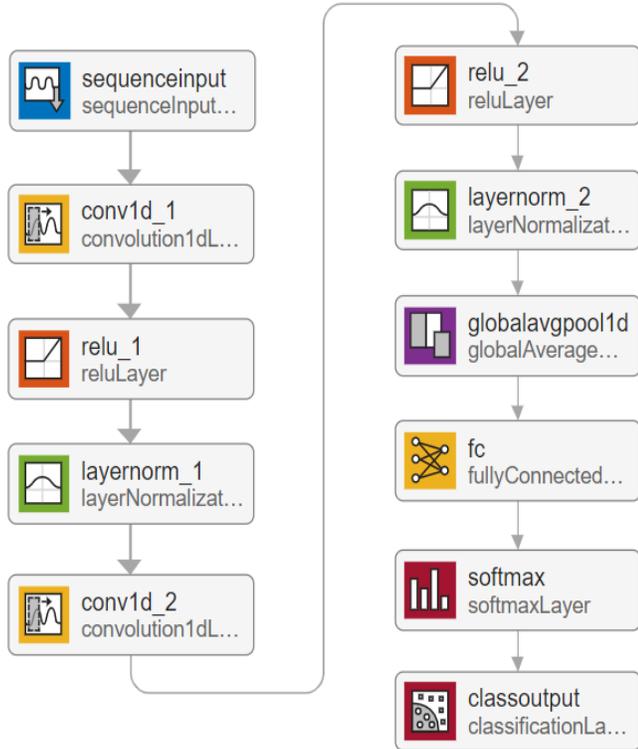


Fig. 8. 1D CNN architecture.

TABLE III. 1D CNN HYPERPARAMETERS

Hyperparameters	Values
Optimizer	adam
No. epochs	300
Learning rate	0.01
Batch size	64
Weights initializer	glorot
Bias ininitializer	zeros
No. layers	11
Input size	34*6
Input size of fully connected layer	128
Output size of fully connected layer	3
No. Classes	3
Validation frequency	500

IV. RESULTS AND DISCUSSION

A. Results

This paper introduces an innovative method for classifying ECG heartbeats into three categories (N, S, and V) following the AAMI standard. The primary goal is to improve accuracy, sensitivity, specificity, and F1 score for each class.

To ensure a fair comparison with existing models, the data was divided into two subsets: DS1 for training and DS2 for validation, aligning with the MIT-BIH arrhythmia database.

Our approach begins by employing the Pan-Tompkins algorithm for QRS detection, enabling the calculation of differences between post-RR intervals and pre-RR intervals.

Subsequently, the WST was utilized to extract morphological features from ECG heartbeats. Combining these features with the calculated post-RR and pre-RR results in a 34x6 matrix for each ECG heartbeat. This matrix is then fed into a specially designed 1D CNN for classification, as illustrated in Fig. 9.

By integrating these techniques, the classification performance aimed to be enhanced and contributed to the field of ECG heartbeat analysis.

After training our 1D CNN model with DS1 heartbeats, it was tested on DS2, revealed notable results. Among 43,478 actual N heartbeats, the model correctly identified 42,891, but intriguingly, it misclassified 191 N heartbeats as V and 396 as S. This misclassification stems from the morphological closeness between N and S classes. Similar challenges were observed in the S class, where out of 1,110 actual S heartbeats, 863 were accurately detected, but 232 were misclassified as N and 15 as V.

Conversely, for the V class heartbeats, out of 3,681 actual V heartbeats, the model successfully classified 3,580, with only 91 misclassified as N and 10 as S. This outcome aligns with the distinct morphological differences observed between V and S.

The precision analysis unveils that the model excels in detecting N and V classes but demonstrates lower precision for S heartbeats. This nuanced understanding is crucial, emphasizing the challenges posed by morphological similarities, especially between N and S classes.

The testing data's confusion matrix, depicted in Fig. 10, encapsulates these findings, providing a visual representation of the model's performance across the three heartbeat categories.

Our model's performance was evaluated using various metrics. Based on the confusion matrix in Fig. 10, our model achieved an average accuracy of 98.71% across three classes. The average precision stood at 87.27%, with an average sensitivity of 91.22%. Additionally, an average specificity of 97.31% and an average F1 score of 89.13% was attained.

B. Discussion

In this research, a comparison with state-of-art models was conducted, utilizing the MIT-BIH arrhythmia database and adopting the same inter-patient data partitioning (DS1 for training and DS2 for testing). The results, illustrated in Table IV, underscore the superior performance of our proposed model across various evaluation metrics.

Notably, for the N class, our model demonstrated remarkable precision at 99.25%, coupled with a specificity of

93.26% and F1 score of 98.95%. Furthermore, our model showcased superiority in the S class, achieving a noteworthy sensitivity of 77.75%. This trend persisted in the V class, where our model outperformed the state-of-art models across multiple metrics, attaining a sensitivity of 97.26%, specificity of 99.54%, and an F1 score of 95.89%. These results underscore the efficacy of our model in effectively distinguishing between N, S, and V.

TABLE IV. COMPARISON WITH STATE OF ART MODELS

Methods	N				S				V			
	SEN	PRE	SPE	F1	SEN	PRE	SPE	F1	SEN	PRE	SPE	F1
Sellami et al. [18]	94	98	82.55	95.95	61.96	52.96	97.89	57.10	87.34	59.44	95.91	70.73
Li et al. [19]	88.52	98.80	91.3	93.37	82.04	30.44	92.8	44.40	92.05	72.13	97.54	80.88
Garcia et al. [20]	94.54	93.33	80.8	93.93	35.22	65.88	98.83	45.90	88.35	79.86	94.92	83.89
Wang et al. [21]	99.04	98.64	87.95	98.83	70.75	77.0	99.51	73.74	94.35	95.32	99.54	94.83
Takalo et al. [22]	91.89	97	76.93	94.37	62.49	55.86	98.11	58.98	89.23	50.58	94.02	64.56
Junaid et al. [23]	98.48	97.39	76.82	97.93	44.01	64.50	99.01	52.32	92.96	89.62	99.22	91.25
Present Work	98.65	99.25	93.26	98.95	77.75	68.01	99.14	72.55	97.26	94.56	99.54	95.89

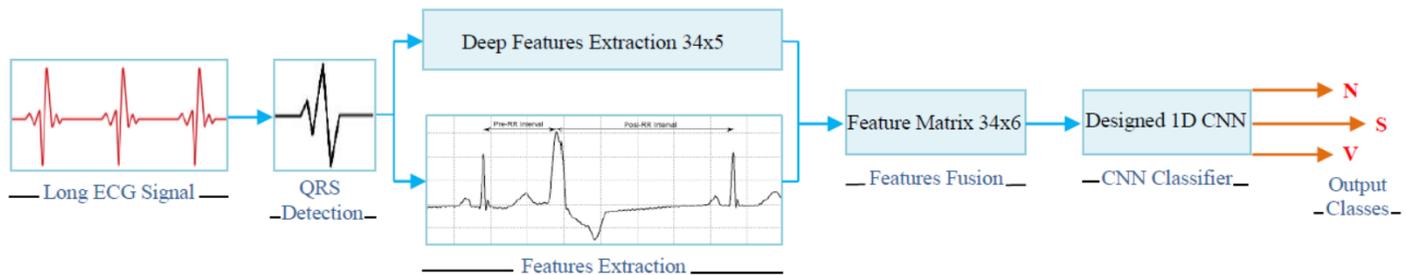


Fig. 9. An overview of the sequential steps for classifying heartbeats

True Class	N	42891	396	191
	S	232	863	15
	V	91	10	3580
		N	S	V
		Predicted Class		

Fig. 10. Testing confusion matrix.

In our investigation, our analysis was extended by comparing the overall accuracy achieved in this study with results from previous works, as detailed in Table V. Sellami et al. [18] achieved an overall accuracy of 92.4% in identifying ECG heartbeats with their proposed approach. In contrast, Li et al. [19] attained an overall accuracy of 88.34% with their method. Among all the other previous works, Wang et al. [21]

achieved the highest overall accuracy of 97.68%. The findings reveal that our proposed model successfully detected 47,334 out of 48,269 instances during testing, yielding an impressive overall accuracy of 98.06%. This value signifies that our model accurately classified 98.06% of the tested heartbeats. Importantly, this overall accuracy underlines the superiority of our model over existing automated classification models.

TABLE V. OVERALL ACCURACY COMPARISON WITH OTHER PREVIOUS WORKS.

Methods	Overall accuracy %
Sellami et al. [18]	92.4
Li et al. [19]	88.34
Garcia et al. [20]	88.99
Wang et al. [21]	97.68
Takalo et al. [22]	89.91
Junaid et al. [23]	95.99
Present Work	98.06

V. CONCLUSION

In this study, an approach was introduced by combining WST for deep feature extraction with 1D CNN classifier. Our model's evaluation, employing inter-patient partitioning, enhances its generalizability for potential clinical applications.

Our results showcase the superiority of our model over state-of-art models. Achieving an overall accuracy of 98.06%, our model excelled in distinguishing between three heartbeat classes. Specifically, it demonstrated a sensitivity of 98.65%, precision of 99.25%, and specificity of 93.26% for the N class. For the S class, our model achieved a sensitivity of 77.75%. Additionally, it outperformed in the V class with a sensitivity of 97.26%, precision of 94.56%, specificity of 99.54%, and an F1 score of 95.89%. These results highlight the significant advancement our model brings to the field.

VI. LIMITATIONS AND FUTURE WORK

Despite these successes, our method has limitations. The computational cost of our feature extraction method is notable due to numerous arithmetic operations during convolution. Furthermore, our model shows room for improvement in evaluating the S class, with an F1 score not exceeding 72.55%.

Future work will concentrate on mitigating computational costs by identifying essential wavelets for optimal results. Additionally, we aim to conduct an in-depth analysis of ECG heartbeats to enhance our model's predictive capabilities for the S class. These advancements will further solidify the applicability and efficiency of our proposed method.

DATA AVAILABILITY

ECG readings were taken from: <https://www.physionet.org/content/mitdb/1.0.0/>

CONFLICT OF INTEREST

We confirm that all authors declare no conflicts of interest.

ACKNOWLEDGMENT

This work was supported by the National Scientific and Technical Research Centre (CNRST), Morocco.

REFERENCES

- [1] M. Elgendi, "Less is more in biosignal analysis: Compressed data could open the door to faster and better diagnosis," *Diseases*, 2018.
- [2] A. Baldassarre, N. Mucci, M. Padovan, A. Pellitteri, S. Viscera, L. I. Lecca, & G. Arcangeli, "The role of electrocardiography in occupational medicine, from Einthoven's invention to the digital era of wearable devices," *International Journal of Environmental Research and Public Health*, 17(14), 4975, 2020.
- [3] W. B. Fye, "A History of the origin, evolution, and impact of electrocardiography," *Am. J. Cardiol.* 73, 937-949, 1994.
- [4] J. R. Henson, "Descartes and the ECG lettering series," *J. Hist. Med. Allied Sci.* 26, 181-186, 1971.
- [5] A. Burguera, "Fast QRS detection and ECG compression based on signal structural analysis," 2019.
- [6] A. S. Abdulbaqi, A. J. Obaid, M. H. Abdulameer, "Smartphone-based ECG signals encryption for transmission and analyzing via IoMTs," *Journal of Discrete Mathematical Sciences and Cryptography*, 24(7), 1979-1988, 2021.

- [7] 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.
- [8] C. A. Martin, G. K. Matthews, C. L. Huang, "Sudden cardiac death and inherited channelopathy: the basic electrophysiology of the myocyte and myocardium in ion channel disease," *Heart* 98, 536-53, 2012.
- [9] D. Da Costa, W. J. Brady, J. Edhouse, "Bradycardias and atrioventricular conduction block," *BMJ, Br Med J* 324, 535-8, 2002.
- [10] M. A. Serhani, H. T. El Kassabi, H. Ismail, A. Nujum Navaz, "ECG monitoring systems: Review, architecture, processes, and key challenges," *Sensors*, 20(6), 1796, 2020.
- [11] S. Hong, W. Zhang, C. Sun, Y. Zhou, H. Li, "Practical lessons on 12-lead ECG classification: meta-analysis of methods from physionet/computing in cardiology challenge 2020," *Frontiers in Physiology*, 12, 2505, 2022.
- [12] H. Wang, Y. Zhou, B. Zhou, X. Niu, H. Zhang, Z. Wang, "Interactive ECG annotation: An artificial intelligence method for smart ECG manipulation," *Information Sciences*, 581, 42-59, 2021.
- [13] M. Sharma, J. S. Rajput, R. S. Tan, U. R. Acharya, "Automated detection of hypertension using physiological signals: A review," *International Journal of Environmental Research and Public Health*, 18(11), 5838, 2021.
- [14] T. Anbalagan, M. K. Nath, D. Vijayalakshmi, A. Anbalagan, "Analysis of various techniques for ECG Signal in Healthcare, Past, Present, and Future," *Biomedical Engineering Advances*, 100089, 2023.
- [15] W. Zeng, B. Su, Y. Chen, C. Yuan, "Arrhythmia detection using TQWT, CEEMD and deep CNN-LSTM neural networks with ECG signals" *Multimedia Tools and Applications*, 82(19), 29913-29941, 2023.
- [16] M. Wasimuddin, K. Elleithy, A. S. Abuzneid, M. Faezipour, O. Abuzaghlh, "Stages-based ECG signal analysis from traditional signal processing to machine learning approaches: A survey," *IEEE Access*, 8, 177782-177803, 2020.
- [17] C. K. Roopa, B. S. Harish, "A survey on various machine learning approaches for ECG analysis," *International Journal of Computer Applications*, 163(9), 25-33, 2017.
- [18] M. E. A. Bourkha, A. Hatim, D. Nasir, S. E. Beid, "Enhanced Atrial Fibrillation Detection-based Wavelet Scattering Transform with Time Window Selection and Neural Network Integration," *International Journal of Advanced Computer Science and Applications(IJACSA)*, 14(12), 2023. <http://dx.doi.org/10.14569/IJACSA.2023.0141252>
- [19] Association for the Advancement of Medical Instrumentation. (1994). American national standard for ambulatory electrocardiographs, publication ANSI. AAMI EC38-1994.
- [20] A. Sellami, H. Hwang, "A robust deep convolutional neural network with batch-weighted loss for heartbeat classification," *Expert Systems with Applications*, 122, 75-84, 2019.
- [21] Y. Li, R. Qian, K. Li, "Inter-patient arrhythmia classification with improved deep residual convolutional neural network," *Computer Methods and Programs in Biomedicine*, 214, 106582, 2022.
- [22] G. Garcia, G. Moreira, D. Menotti, E. Luz, "Inter-patient ECG heartbeat classification with temporal VCG optimized by PSO," *Scientific reports*, 7(1), 10543, 2017.
- [23] T. Wang, C. Lu, Y. Sun, M. Yang, C. Liu, C. Ou, "Automatic ECG classification using continuous wavelet transform and convolutional neural network," *Entropy*, 23(1), 119, 2021.
- [24] J. Takalo-Mattila, J. Kiljander, J. P. Soininen, "Inter-patient ECG classification using deep convolutional neural networks," In 2018 21st Euromicro Conference on Digital System Design (DSD) (pp. 421-425). IEEE, August 2018.
- [25] J. Malik, O. C. Devecioglu, S. Kiranyaz, T. Ince, M. Gabbouj, "Real-time patient-specific ECG classification by 1D self-operational neural networks," *IEEE Transactions on Biomedical Engineering*, 69(5), 1788-1801, 2021.
- [26] Z. He, Y. Chen, S. Yuan, J. Zhao, Z. Yuan, K. Polat, ... A. Hamid, "A novel unsupervised domain adaptation framework based on graph convolutional network and multi-level feature alignment for inter-subject ECG classification," *Expert Systems with Applications*, 221, 119711, 2023.

- [27] G. B. Moody, R. G. Mark, "The impact of the MIT-BIH arrhythmia database," *IEEE engineering in medicine and biology magazine* 20(3), 45-50, 2001.
- [28] J. Pan, W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE transactions on biomedical engineering*, (3), 230-236, 1985.
- [29] M. A. Belkadi, A. Daamouche, "A robust QRS detection approach using stationary wavelet transform," *Multimedia Tools and Applications*, 80(15), 22843-22864, 2021.
- [30] S. Talukder, R. Singh, S. Bora, R. Paily, "An efficient architecture for QRS detection in FPGA using integer Haar wavelet transform," *Circuits, Systems, and Signal Processing*, 39, 3610-3625, 2020.
- [31] W. Cai, D. Hu, "QRS complex detection using novel deep learning neural networks," *IEEE Access*, 8, 97082-97089, 2020.
- [32] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, 16, 321-357, 2002.
- [33] J. Andén, S. Mallat, "Deep scattering spectrum," *IEEE Transactions on Signal Processing*, 62(16), 4114-4128, 2014.