Kufa Journal of Engineering
Vol. 16, No. 2, April 2025, P.P. 423 -444
Article history: Received 5 October 2024, last revised 20 December 2024, accepted 20 December 2024

A NOVEL DEEP 2D-CNN MODEL FOR ECG-BASED ARRHYTHMIA DIAGNOSIS WITH SELECTIVE ATTENTION MECHANISM AND CWT INTEGRATION

Hassanain Shakir Mansour[1,2], Morteza Valizadeh[3], Alaa Hussein Abdulaal[4] and Mehdi Chehl Amirani[5]

[1] Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, West Azerbaijan, Iran.

[2]Department of Computer Technical Engineering, College of Information Technology, Imam Ja'afar Al-Sadiq University, Al-Muthanna, Iraq.
Email: Hasanain.shakir@ijsu.edu.iq

[3] Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, West Azerbaijan, Iran. Email: Mo.valizadeh@urmia.ac.ir

[4] Department of Electrical Engineering, College of Engineering, Al-Iraqia University, Baghdad, Iraq. Email: Engineeralaahussein@gmail.com

[5] Department of Electrical Engineering, Faculty of Electrical and Computer Engineering, Urmia University, West Azerbaijan, Iran. Email: M.amirani@urmia.ac.ir .

https://doi.org/10.30572/2018/KJE/160225

ABSTRACT

This study introduces an innovative approach for arrhythmia diagnosis via electrocardiogram (ECG) signals, employing a 2D Convolutional Neural Network (CNN) model fused with a Continuous Wavelet Transform (CWT) and a Selective Attention Mechanism (SAM). The SAM enhances feature focus, improving classification accuracy. The model effectively categorizes ECG signals into Normal and Abnormal classes, subcategorizing Abnormal patterns into four types. Merging deep learning with signal processing enables the correct classification of arrhythmia. The outcome yields high accuracy, such as 99.784% for multi-class and 99.94% for binary classification. The methodology proposed in this paper shows the capability of secure yielding for arrhythmia diagnosis, setting novel standards within healthcare applications.

KEYWORDS

Arrhythmia, ECG signals, CNN, Continuous wavelet transform, Selective attention mechanism.

1. INTRODUCTION

Cardiovascular diseases (CVDs) are a leading cause of death worldwide (Gaidai et al., 2023), mandating advanced diagnostic tools and techniques in the fight against these diseases. Electrocardiography has formed the cornerstone of cardiac health assessment for many years and provides extremely useful details about the heart's electrical activity (Ardeti et al., 2023), (Talib et al., 2020). For the myriad cardiac disorders, arrhythmias pose a challenge in clinical practice because of the wide range of their manifestations and possible complications (Giancaterino et al., 2020).

Arrhythmias are group terms that refer to many abnormal heart rhythms and are, therefore, one of the most common cardiac disorders, but with varied clinical implications (Liu et al., 2023). Such disturbances in the heart's electrical activity express themselves as tachycardias, bradycardias, or irregular heartbeats, significantly interfering with the normal rhythm and coordination of cardiac contractions. An arrhythmia may result from many causes: structural heart disease, electrolyte disturbances, genetic predispositions, or systemic diseases like hypertension and diabetes. Thus, classifying and diagnosing an arrhythmia in a clinical setting is essential for properly managing and treating the patient. Conventional arrhythmia detection methods involve a professional's visual inspection and interpretation of ECG waveforms to recognize characteristics that define specific arrhythmias. Manual analysis is time-consuming, subjective, and liable to human error, mostly in subtle or complex arrhythmias.

Traditional arrhythmia detection has been done by the clinician observing the electrocardiograms manually, which, although expert-driven, remains prone to human error and intervariability; hence, the need for more objective methods.

Within the last ten years, machine learning models have grown to be among the most powerful methods for automatic arrhythmia classification (Mathur et al., 2020). Other approaches, such as support vector machines (Abdulaal et al., 2024d) and random forests (Myrovali et al., 2022), dig even deeper into the statistical patterns in ECG data. While such models do relatively well, they often require careful feature selection and tuning, which costs scalability and, hence, generalization to diverse datasets.

This gets further advanced with the introduction of deep learning architectures (Mohammed et al., 2022), especially those involving CNNs that allow end-to-end learning from the raw ECG signals (Ahmed et al., 2023), which are altering the medical scenario related to the diagnosis and detection of arrhythmia, QRS wave, and ST segment, as shown in Fig. 1.

This approach thus allows researchers and clinicians to increase the accuracy and efficiency of arrhythmia classification since CNN can automatically extract salient features from the ECG signal. Since CNN can capture spatial dependencies in data, it would be well-positioned to detect minute deviations and patterns within the waveform of ECG, which may elude human perception. Furthermore, using CNNs in arrhythmia analysis will further the development of automated systems capable of real-time arrhythmia detection, thus providing the opportunity for quick, homogeneous, and objective assessments (Ullah et al., 2021). Being based on machine learning, it will enable healthcare providers to enhance diagnostic potential, improve patient outcomes, and probably even uncover novel insights into the mechanisms underlying arrhythmias.

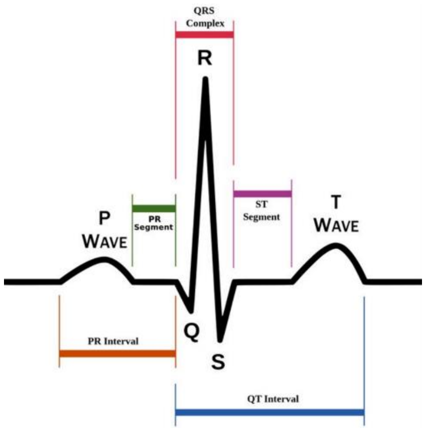


Fig. 1. ECG signal

Integrating CNN-based methods into arrhythmia diagnosis is one bright frontier of cardiovascular health care, foreshadowing a new age of precision medicine and individualized treatment strategies. As investigators refine and expand machine learning applications in arrhythmia analysis, this field is on the verge of substantial improvement in diagnosing, monitoring, and managing arrhythmic disorders, making it possible to enhance the quality of life for patients and health systems.

Over many years, numerous researchers have dedicated considerable efforts to automating the detection of cardiac arrhythmias. A common choice for this research has been the MIT-BIH arrhythmia database (Moody et al., 2001), a widely available resource in this field. Different researchers have developed various models to classify ECG signals. Despite the use of diverse datasets by these researchers, the categorization is primarily based on the MIT-BIH database, focusing on the classification of the same heartbeats, which enables a meaningful comparison across studies.

In previous studies on diagnosing arrhythmias within the MIT-BIH database, researchers (Yıldırım et al., 2018), (Acharya et al., 2017) utilized deep genetic hybrid classifiers on extended ECG data, achieving an accuracy of 94.6%. Given ECG signals' nonlinear and dynamic nature in real-world scenarios, higher-order statistical (HOS) methods, exemplified in

(Martis et al., 2013), can effectively capture subtle variations. By employing principal component analysis to reduce the dimensionality of spectrum features, researchers enhanced automatic pattern recognition using a least squares-support vector machine and a four-layer feed-forward neural network, achieving an impressive average accuracy of 93.48%.

Despite the common trend of declining accuracy with increasing network depth in deep neural network training, a notable exception was found in (Li et al., 2017), where a basic CNN model with five layers attained a remarkable accuracy of 97.5%. The method followed was to use wavelet transform-based quadratic waves to extract individual ECG waveforms, after which a fiduciary marker array could be used for classification through a probabilistic neural network, achieving an accuracy rate of 92.7% (Gutiérrez-Gnecchi et al., 2017). Other methods have used the discrete wavelet transform for denoising signals.

(Chen et al., 2020) classified six different types of arrhythmias using convolutional neural networks (CNNs) and long short-term memory networks (LSTM) models. This model adds a noise removal process to the ECG signal's pre-processing stage and feeds this denoised signal as input to the CNN architecture. The SoftMax layer forms the last layer of the model, which generates four varied outputs corresponding to the six types of arrhythmias.

(Ullah et al., 2020) proposed a new technique using a two-dimensional Convolutional Neural Networks model for classifying ECG signals in arrhythmia diagnosis. That is the essence of deep learning: solving problems created by high-dimensional data without explicit rules or patterns.

Such models depict superior performance by automatically extracting complicated features with no human interference. However, such models also bring many other challenges, such as high computation requirements and reduced interpretability of the results, which are key issues for clinical applications.

Despite this, there are still certain lacunae in the classification of arrhythmias. Clinical applications, on a larger scale, are restricted due to limitations in the accuracy, computational needs, and partial interpretability of the models. These lacunae need to be identified to develop robust and reliable diagnostic tools.

While the present study addresses these challenges, it presents a new approach in arrhythmia classification to improve accuracy and interpretability while reducing computational burdens. This paper presents a sophisticated approach involving a 2D CNN model, improved with Continuous Wavelet Transform and a selective attention mechanism for analyzing ECG signals, which can further classify them into two broad classes: Normal and Abnormal. It can, therefore, further differentiate four subtypes under the abnormal category as follows: Ventricular

Premature Contraction (VPC), Left Bundle Branch Block (LBB), Atrial Premature Contraction (APC), and Right Bundle Branch Block (RBB), as shown in Fig. 2.

VPC is the type of premature heart contraction that occurs from the ventricles, mostly skipping a beat. LBB is a block or a delay in the impulses along the left bundle branch; it delays the left ventricle contraction of the heart. APC is among the basic beats that happen extra from the atria and can create irregular heartbeats. RBB is a blockage or retardation of the normal electrical impulses to the right bundle branch, which delays the contraction of the right ventricle.

This comprehensive framework utilizes deep learning and signal processing techniques to realize accurate and detailed classification for different arrhythmia patterns in ECG data.

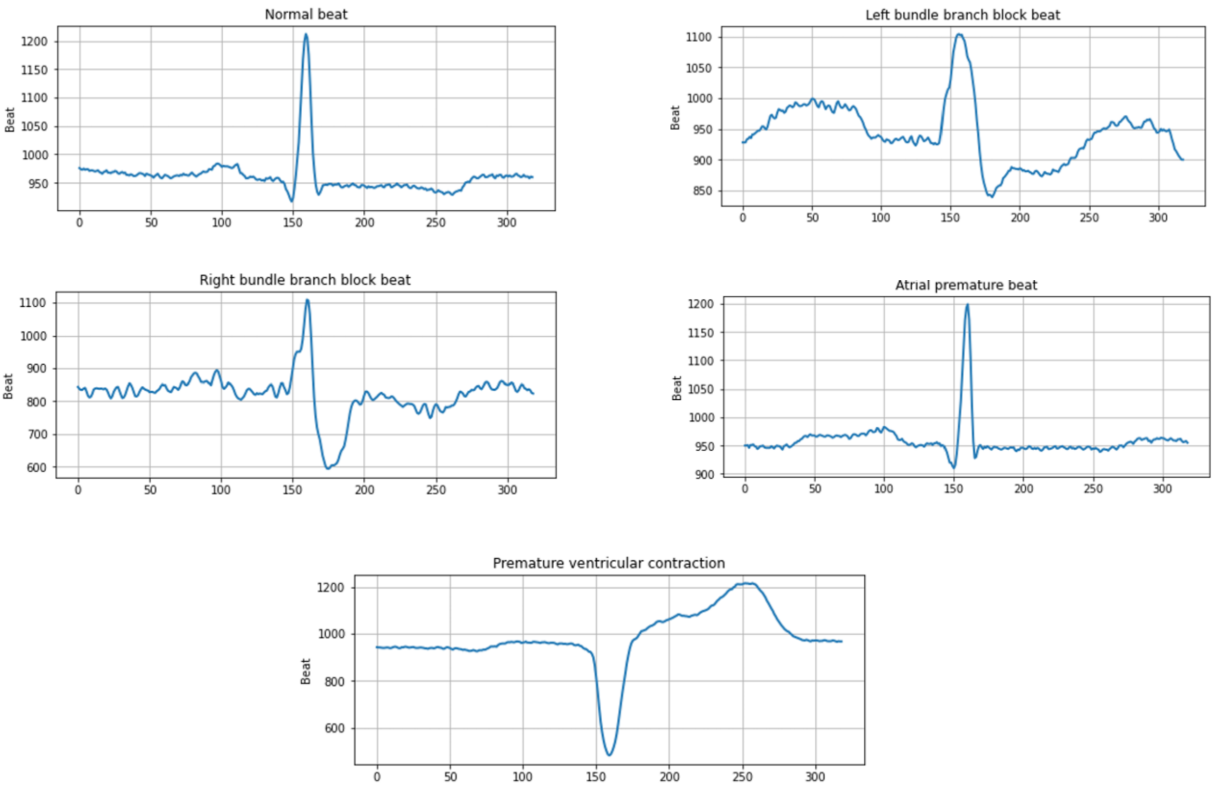


Fig. 2. MIT-BIH Arrhythmia heartbeats (Cheikhrouhou et al., 2021)

2. MATERIALS AND METHODS

2.1. MIT-BIH dataset

The MIT-BIH dataset (Moody et al., 2001) comprises 48 segments of 2-channel ambulatory ECG records, each lasting thirty minutes, obtained from 47 patients analyzed by the BIH Arrhythmia Laboratory. Among these, twenty-three recordings were randomly selected from a pool of 4000 24-hour ambulatory ECG recordings originating from a diverse mix of indoor (60 out of 100) and outdoor patients (40 out of 100). The remaining twenty-five recordings were explicitly chosen from a similar dataset to include significant yet clinically meaningful arrhythmias that might not be well represented in a less complex sample (Ullah et al., 2021a). These recordings were digitized at 360 samples per second per channel with an 11-bit resolution spanning a range of 10 mV. Each record underwent independent review by more than two cardiologists, and any discrepancies were resolved to generate computer-readable diagnostic annotations for each record, totaling approximately 110,000 annotations in this database. The dataset encompasses various abnormal heart rhythms, such as left and right bundle branch blocks, atrial premature contractions, and ventricular premature contractions, representing a range of heart block conditions that can manifest.

2.2. Arrhythmia Classification Method

In an innovative approach to boosting the accuracy of arrhythmia classification, traditional 1D ECG signals are transformed into informative 2D ECG images. These images are input into the task-specific 2D CNN model (Ullah et al., 2021a). Hence, the methodology has significant signal processing steps, such as generating 2D images, data augmentation, feature extraction, and data categorization. A selective attention mechanism is desirably designed in combination with converting 1D ECG signals into insightful 2D ECG images, all to enhance the accuracy of arrhythmia classification. After transformation, these images are fed into a state-of-the-art 2D CNN model with a selective attention mechanism focusing on key regions in ECG images. The very detailed architecture of the advanced model is depicted in Fig. 3. This can be observed in the structured steps followed in methodology: initial signal processing, dynamic 2D image generation, robust data augmentation, feature extraction with selective attention, and precise data categorization.

2.3. Data Preprocessing

The preprocessing step was essential to change these preliminary 1D ECG signals into complex 2D images to be fed into the 2D CNN model to extract various features embedded in the data. This helps in effectively classifying the different ECG patterns. ECG patterns extracted from the MIT-BIH database are precisely separated into individual ECG beats by using the Q-wave peak time and further annotated according to the type of arrhythmia under question.

This method truncates the previous and next 20 ECG waves around the Q-wave peak signals in mapping 2D ECG images in line with Q-wave peak timing information (Ullah et al., 2021a). Enabling the marking of the extent for every ECG wave organizes an accurate marking of the ECG wave intervals

2.4. Continuous wavelet transforms (CWT)

Traditionally, 1D convolutional neural networks have been the backbone of processing 1D signals; these networks have been less flexible. Given these limitations, there has been a paradigm shift in using 2D representations (Mohonta et al., 2022). The two-dimensional CNN models further enhance flexibility and effectiveness in handling complex temporal patterns through 2D kernels, hence mitigating some of the errors that their 1D versions could have carried forward. This transition unleashes a host of powerful techniques suited explicitly for nuanced portrayal, firmly realizing 2D data structures and enabling in-depth exploration of the information underlying them.

2D CNN architectures are used to capture microstructural nuances with a high degree of meticulousness that is embedded within input data. More specifically, electrocardiograms are attained by transforming the latter into 2D spectrograms using the Continuous Wavelet Transform methodology (1). It delineates non-stationary features that have different instantaneous frequencies and gives the ability to encapsulate the intricate details existing within the frequency domain of the continuous wavelet transform (CWT).

The CWT methodology embeds information about the frequency dynamics and the localized amplitudes of time-varying signals (2), fully expressing the underlying data characteristics. A modicum of signal stationarity will be maintained within finite temporal windows to model this relationship. These expressions capture this transformation from the original 1D signal to richly detailed 2D images (512 × 512, in the case of Eq. 3 and 4), bringing out this transformation's heart.

$$X[a,b] = \frac{1}{c\psi}\int_{-\infty}^{+\infty} x(t)\psi_{a,b}^{*}(t)dt \tag{1}$$

Where X[a,b] represents the spectrogram

$$Ximage(i,j) = X\left(\frac{i}{512}xT, \frac{j}{512}xF\right) \tag{2}$$

Where i,j range from 1 to 512, and T, F represent the spectrogram's total time and frequency span, respectively. To ensure that the spectrogram values are within the range [0, 255] for image representation, the values can be normalized:

$$Xnormalized(i,j) = \frac{255x\,(Ximage(i,j) - \min(Ximage))}{\max(Ximage) - \min(Ximage)} \tag{3}$$

The frequency domain representation via Continuous Wavelet Transform (CWT):

$$CWT_x(a,b) = \int_{-\infty}^{+\infty} x(t)\frac{1}{\sqrt{a}}\psi^{*}\left(\frac{t-b}{a}\right)(t)dt \tag{4}$$

These formulas allow for the transformation of spectrogram data into image data suitable for visualization in a 512 × 512 format.

2.5. Data Augmentation

Data augmentation involves enriching the existing dataset by incorporating supplementary relevant information, reducing the need for manual data expansion, and potentially enhancing data quality. This process is precious when using 2D images as inputs for CNN models, as it can improve model generalization and robustness by providing a more diverse set of training examples (Ullah et al., 2021a).

The specific transformations applied in this work on ECG images included rotation, scaling, and flipping. Rotation introduces the appearance of ECG signals from different orientations, making the model invariant toward rotations, hence generalizing the learning better. Scaling concerns the size of an image and helps a model to recognize ECG patterns of any size, which means it will make the model more robust against the scale change. Moreover, flipping-horizontal or vertical-introduces variation to prevent the model from being biased toward any one orientation.

All these steps increase the variability of the training data, thereby allowing the generalization of new unseen data. All these augmentations are supposed to enhance robustness, as that would avoid overfitting of the model and allow learning more invariant features. While data augmentation does not guarantee any improvement, this work tries to maximize such a gain by choosing transformations that best align with the characteristics of ECG data.

3. PROPOSED MODEL

3.1. Architecture of 2D CNN with Selective Attention Mechanism (SAM)

Innovating upon the conventional 2D CNN architecture for ECG classification, this advanced model incorporates a selective attention mechanism to augment its diagnostic capabilities, as shown in Table 1 and Fig. 3. The intricate nature of 2D image data poses challenges for standard feed-forward networks due to the parameter surge. However, integrating diverse nonlinear filters within CNNs enables the extraction of fine-grained local features by correlating neighboring pixels.

This study pioneers transforming 1D ECG signals into 2D ECG images by leveraging convolutional and down-sampling layers, enhancing classification accuracy to levels comparable to expert diagnosis. Adding a selective attention mechanism within the architecture further refines the model's ability to focus on salient features critical for ECG analysis. This enhancement aims to revolutionize the traditional approach to ECG arrhythmia diagnosis, potentially surpassing the diagnostic acumen of human experts. The model's architecture, shown in Fig. 3, is a sophisticated, artful weave of layers designed for optimized feature extraction and accuracy in classification, which shows how sophisticated the model is and its potential for transformative impact in medical diagnosis.

Table 1. 2D-CNN model layers

Layer	Type	Kernel Size	Stride	Kernel	Input Size
1	2D Conv 1	3 × 3	1	64	512 × 512 × 1
2	Pooling 1	2 × 2	2	-	512 × 512 × 64
3	2D Conv 2	3 × 3	1	128	256 × 256 × 64
4	Pooling 2	2 × 2	2	-	256 × 256 × 128
5	2D Conv 3	3 × 3	1	256	128 × 128 × 128
6	Pooling 3	2 × 2	2	-	128 × 128 × 256
7	2D Conv 4	3 × 3	1	512	64 × 64 × 256
8	Pooling 4	2 × 2	2	-	64 × 64 × 512
9	Selective Attn	-	-	-	32 × 32 × 512
10	New_FC			4096	16 × 16 × 512
11	New_Output			5	4096

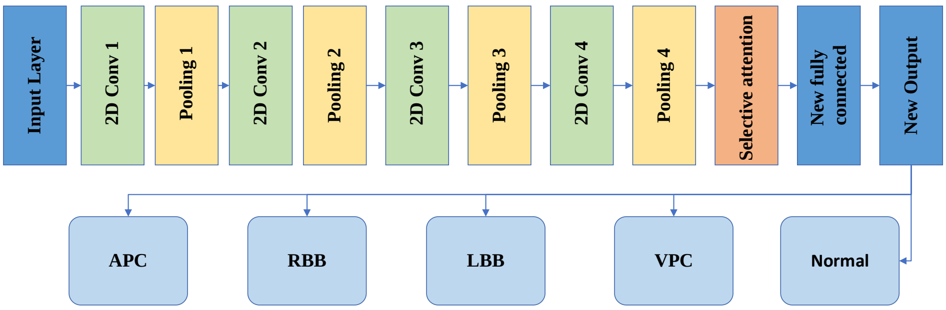


Fig. 3. The architecture of the 2D-CNN model

A four-layer convolution architecture was the most tempting for the model to learn the complex pattern in ECG, ensuring the presence of hierarchical feature extraction from low-level details to high-level representative key items for an accurate diagnosis. Furthermore, this model combines the advantage of Continuous Wavelet Transform with a time-frequency analysis for a selective attention mechanism that underlines the most relevant features, hence improving the overall accuracy and robustness. A deeper network ensures the attention mechanism works on rich high-dimensional feature maps.

3.2. Selective Attention Mechanism (SAM)

SAM enables attention to be focused on specific regions or characteristics in the input data. Drawing from the approach in ECG signal analysis (Wang et al., 2023), a SAM might draw attention to some patterns or segments in a signal that are crucial for correct discrimination among various arrhythmias. This selective-attention approach first segments the ECG signal into smaller segments (5) and then computes attention scores concerning these segments (6), (7). These scores indicate how important and relevant each segment is concerning arrhythmia diagnosis. For those segments that can be considered more informative and hence carry greater weight in the classification process, a higher value of attention score (8) is given. Each segment is multiplied by its respective attention score (9), giving more importance to segments with higher scores. This can be incorporated into an ECG signal classifier by adding a specifically designed layer called "SelectiveAttentionLayer" to the model architecture (10). This layer works on the ECG signal segmentation principle and then computes attention scores for those segments.

$$X = \{x_1, x_2, x_3, x_4, \ldots, x_n\} \tag{5}$$

Where X (represents the ECG segment) is divided into n segments

$$a_i = \{a_1, a_2, a_3, \ldots.., a_n\} \tag{6}$$

a_i is the attention score for segment X_i

$$a_i = \frac{\exp(e_i)}{\sum_{j=1}^{n} \exp(e_j)} \tag{7}$$

Where e_i represents the raw score for segment X_i, calculated based on its features, this normalization ensures that attention scores a_i sum to 1, making them comparable across segments.

$$X_{Max} = \arg Max_i \{a_i\} \tag{8}$$

X_{Max} represents the segment with the highest attention score, indicating which part of the ECG signal is most informative for arrhythmia detection.

$$X_{weighted} = \sum_{i=1}^{n} a_i X_i \tag{9}$$

$X_{weighted}$ is a weighted sum of all segments, multiplying each segment by its respective attention score. It aggregates information, giving more importance to segments with higher scores.

$$SAM(X) = X_{weighted} \tag{10}$$

The class SAM takes a tensor representing the ECG signal as input at prediction time. Intuitively, attention scores would reflect the importance of each segment in the context of arrhythmia classification. This means that the segment with the highest attention score can be considered most informative and prioritized, as shown in Fig. 4. In these key segments, the SAM will extract key temporal patterns related to different arrhythmias in the ECG signal. This will enhance the model's capability to perceive the patterns clearly; hence, the accuracy and reliability of diagnosis in arrhythmia will probably be improved. Attention-based models working selectively on relevant segments of the ECG signals provide valuable insights into the presence and exact types of arrhythmias, thereby allowing timely, well-informed decisions to be made at the clinical level. With growing interest in attention-based methods, this phenomenon is gradual in its approach toward advanced diagnosis of arrhythmias and improved patient outcomes related to cardiovascular health care.

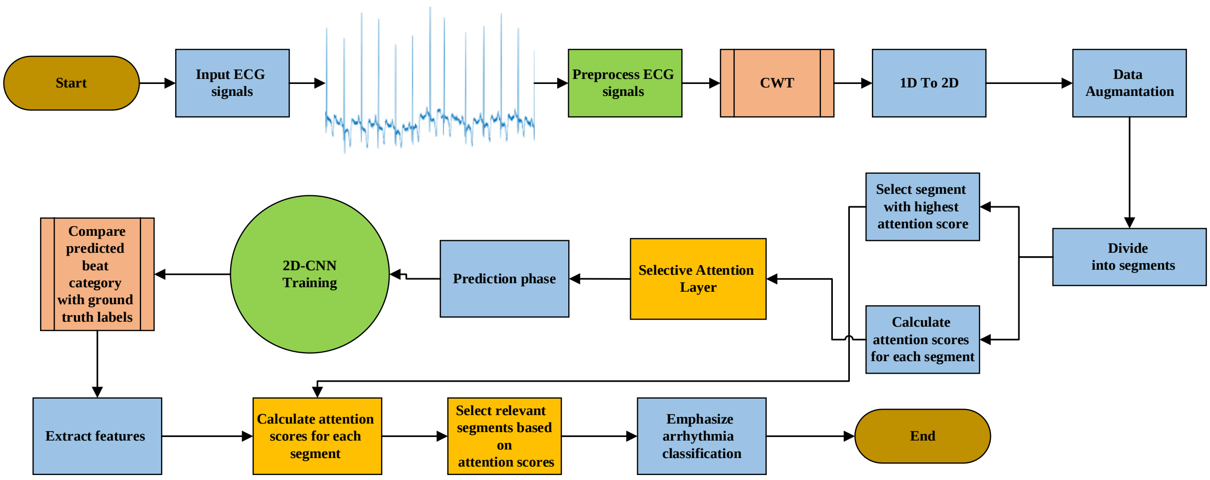


Fig. 4. Flow chart of the proposed model

This further makes model parameter selection the most important decision that needs to be made in order to realize the 2D Convolutional Neural Network architecture with the Selective Attention Mechanism for optimal performance. Kernel size and stride have been used in CNN architecture to capture the spatial hierarchies of 2D ECG signal representations effectively. These parameters were chosen from empirical tests and established practices in related areas.

The integration of CWT for pre-processing transforms 1D ECG signals to 2D image features. This would make it easy to extract the time-frequency features, allowing CNNs to learn complex patterns that intrinsically come with the ECG data.

This selective attention mechanism allows the model to focus on the parts of the input data that contribute most to improving the classification accuracy. Attention layers are chosen based on how they effectively emphasize critical features, as supported by related research.

3.3. Activation Function

The activation function is key in CNN classification models' finger-kernel sizing and output weighting. Recent research has mainly followed non-linear activation functions like the Leaky Rectified Linear Units function (LReLU), the Rectified Linear Units function (ReLU), and the Exponential Linear Units function (ELU) in the finger-kernel sizing and output weighting. ReLU is one of the most common activation functions because of its excellence in setting all negative values to zero. That trait gives it an advantage of sparsity in the network (Ullah et al., 2021a) (Hao et al., 2020). This also provides a way for feature selection.

ReLU became one of the foundations of CNN algorithms because it suppresses negative values in computations. Indeed, activation functions such as LReLU and ELU have low use in classification tasks because they allow small negative values to move forward into the next layer. Much empirical evidence shows that ReLU is better than LReLU in arrhythmia classification applications since it enhances the model's overall performance.

Mathematically, the ReLU activation function (10) can be written as:

$$f(x) = \max(0, x) \tag{10}$$

This relatively simple yet powerful function has become a staple in CNN architectures that perform very well in classification tasks, such as accurately classifying ECG signals in arrhythmia detection.

3.4. 5-fold cross-validation

Implementing 5-fold cross-validation is one of the most important strategies in the robust evaluation of models developed for ECG arrhythmia classification using a CNN model. The dataset containing all ECGs is divided into five subsets where the model is trained on four folds and validated on the fifth fold. The process is repeated five times (Abdulaal et al., 2024c), wherein all ECG signals used would have been applied once for training and another time for validation to gauge the model's performance fully. Such 5-fold cross-validation reduces variance, detects overfitting or underfitting, and improves the predictive accuracy and generalizability of the model. This would provide insight into how effective the model is at picking up arrhythmia patterns, how to optimize hyper-parameters, and how to improve the overall quality of the classification system.

3.5. Training process

The model training process involves several key steps to ensure optimal performance. In other words, the input of the CNN, which is designed to consist of four convolution layers with a selective attention mechanism, is a 2D image representation of the 1D ECG signals obtained using Continuous Wavelet Transform. The selected key hyperparameters are chosen as follows: the learning rate is kept at 0.001 for a balanced convergence speed with stability; batch size is taken as 32, enabling efficiency in training with good generalization for the model; and the number of epochs is put at 50-adequate for sufficient training without leading to overfitting. It applies the Adam optimizer, as it is efficient and maintains adaptive learning rates; the ReLU activation function introduces non-linearity into the model, hence the ability to pick complex patterns. It further improves the capability of giving much attention to important features through the selective attention mechanism, which improves generalization capability. The model will conduct 5-fold cross-validation to make it robust and reliable over subsets of data. This means that during this training, the performance is monitored with a validation set using various techniques to avoid overfitting, such as early stopping so that it reaches a strong and efficient model appropriate for the task.

3.6. Evaluation metrics

The evaluation metrics used in this paper are as follows (Abdulwahhab et al., 2024), (Abdulaal et al., 2024a):

Accuracy (11): The proportion of true results (both true positives and true negatives) among the total number of cases examined (Abdulaal et al., 2024b).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} X\ 100 \qquad (11)$$

Recall (12): The proportion of true positives among the total number of actual positives.

$$Recall = \frac{TP}{TP + FN} X\ 100 \qquad (12)$$

Precision (13): The proportion of true positives among the total number of positive predictions.

$$Precision = \frac{TP}{TP + FP} X\ 100 \qquad (13)$$

Specificity (14): The proportion of true negatives among the total number of actual negatives.

$$Specificity = \frac{TN}{TN + FP} X\ 100 \qquad (14)$$

F1 score (15): The harmonic mean of precision and recall (Abdulaal et al., 2024e).

$$F1-Score = \frac{TP}{TP + \frac{1}{2}(FP + FN)} X\ 100 \qquad (15)$$

The Receiver Operating Characteristic (ROC) curve plots the True Positive Rate (Recall) against the False Positive Rate to evaluate a binary classifier's performance across different threshold values (Janssens et al., 2020).

Where:

TP: True Positives, FP: False Positives, FN: False Negatives, and TN: True Negatives.

4. RESULTS

These study results highlight that, compared to traditional methods, the proposed methodology is further improved for diagnosing arrhythmias using ECG signals. In this sense, the 2D CNN architecture, combined with selective attention, performs optimally in classifying arrhythmias with extremely high detail and precision.

The selective attention mechanism for the 2D CNN framework has significantly enhanced the former procedure of apprehending ECG signals. This new setup makes it possible for attention to be drawn to the major characters of the ECG data, thus enhancing the model in terms of the separation of subtle patterns and aberrations typical of various types of arrhythmias. It achieves the extraction of the spatial relationships intrinsic within ECG signals, represented by the 2D CNN architecture. It concurrently considers the ability of the attention mechanism to concentrate on important information. The proposed model captures complex subtleties of arrhythmic patterns.

The results unequivocally show the model's supremacy for the following remarkable performance metrics: Accuracy: 99.784%, Precision: 99.782%, F1-Score: 99.776%, and Specificity: 99.944%, as represented in Table 2. Such metrics reflect the model's extraordinary accuracy, precision, F1-Score, and specificity in classifying arrhythmias effectively, showcasing its robustness and effectiveness in diagnosing cardiac abnormalities.

Table 2. Evaluation metrics of multi-class classification

CNN	Metrics	N	LBB	RBB	APC	VPC	Average
2D-CNN	Accuracy	97.55%	98.48%	97.14%	96.34%	93.14%	96.53%
	Precision	96.95%	95.08%	95.91%	97.10%	97.73%	96.55%
	Error	2.45%	1.52%	2.86%	3.66%	6.86%	3.47%
	F1-Score	97.25%	96.75%	96.52%	96.72%	95.38%	96.52%
	Specificity	99.23%	98.73%	98.97%	99.28%	99.46%	99.13%
2D-CNN +SAM	Accuracy	99.87%	99.94%	99.90%	99.86%	99.35%	99.784%
	Precision	99.61%	99.78%	99.93%	99.66%	99.93%	99.782%
	Error	0.13%	0.06%	0.1%	0.14%	0.65%	0.216%
	F1-Score	99.74%	99.86%	99.92%	99.76%	99.64%	99.776%
	Specificity	99.90%	99.95%	99.98%	99.91%	99.98%	99.944%

The study introduces a groundbreaking methodology for binary classification between Normal and Abnormal ECG signals, employing a state-of-the-art 2D CNN architecture with a selective attention mechanism. This model leads to an outstanding accurate differentiation between these classes with very impressive performance metrics: Accuracy 99.70%, Precision 99.94%, Recall 99.47%, F1-Score 99.70%, and Specificity 99.94%, as indicated in Table 3. In setting a new benchmark in cardiovascular health research by revolutionizing the analysis of the ECG signal and providing better anomaly detection, this model will bring about much better results in patient outcomes and lead to transformational health practices concerning the diagnosis and management of cardiac conditions.

Table 3. Evaluation metrics for Binary classification

CNN	Accuracy	Precision	Recall	F1-Score	Specificity
2D-CNN	94.91%	94.47%	95.32%	94.89%	94.52%
2D-CNN + SAM	99.70%	99.94%	99.47%	99.70%	99.94%

Fig. 5 shows an attention map representing the most informative features by underlining different grades of importance across the regions. The selective attention mechanism inside this map would enhance attention toward those regions that best describe the underlying regularities or anomalies. The current approach creates opportunities for better interpretive knowledge extraction from the ECG and allows more precise interpretation that might improve diagnostic results by underlining critical signal characteristics.

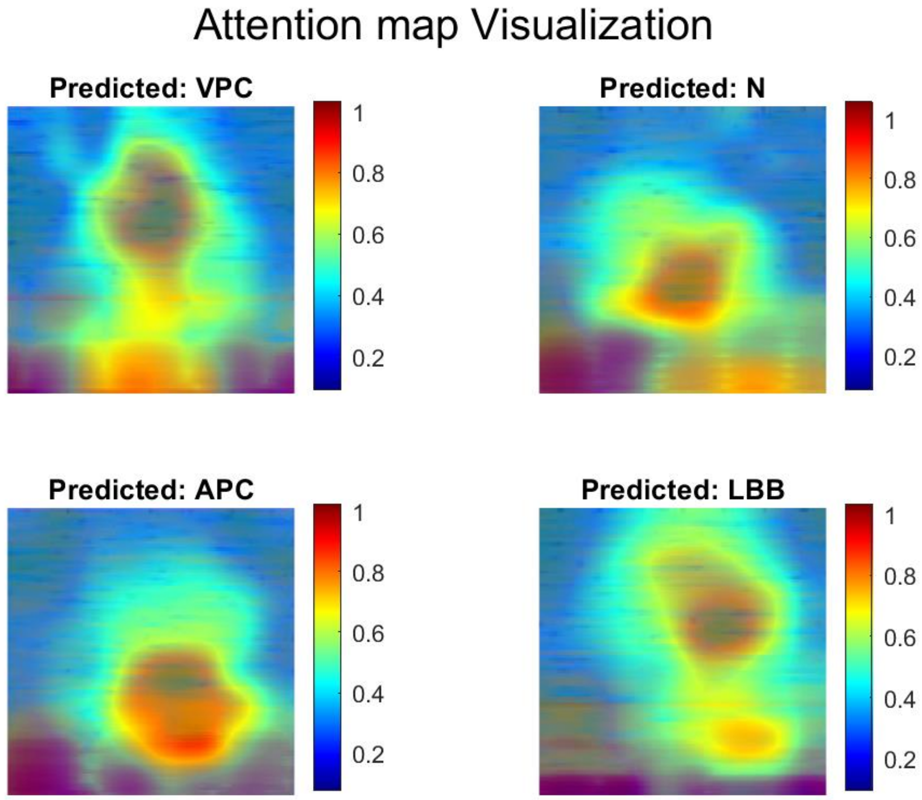


Fig. 5. Attention map visualization

In Fig. 6, the feature extraction process unveils the acquired insights of a neural network tailored for binary classification (normal and abnormal) of arrhythmias derived from ECG signals. Each subplot within this visualization illustrates feature activations about a specific test image, presented as bar plots where the x-axis signifies feature indices and the y-axis denotes activation values. Elevated bars signify features deemed most critical for classifying a given image. The subplot titles in this figure reveal the actual class of the arrhythmia for each signal, allowing for a visual juxtaposition across classes. This lets observers discern which features are pivotal in distinguishing between different arrhythmia classes, providing valuable insights into the model's decision-making process. Such insights could pave the way for enhancements in diagnostic tools and treatment strategies, fostering advancements in healthcare practices.

In a confusion matrix, entries represent the normalized counts of correct and incorrect predictions for each class in multi-class classification, aiding in evaluating model performance across different classes, as shown in Fig. 7. In binary classification, the matrix displays the normalized counts of true positives, false positives, true negatives, and false negatives, providing insights into the model's accuracy and error rates as shown in Fig. 8.

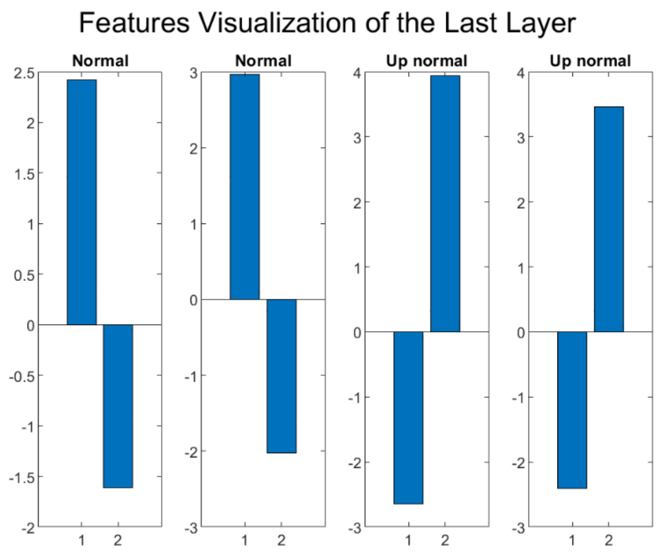
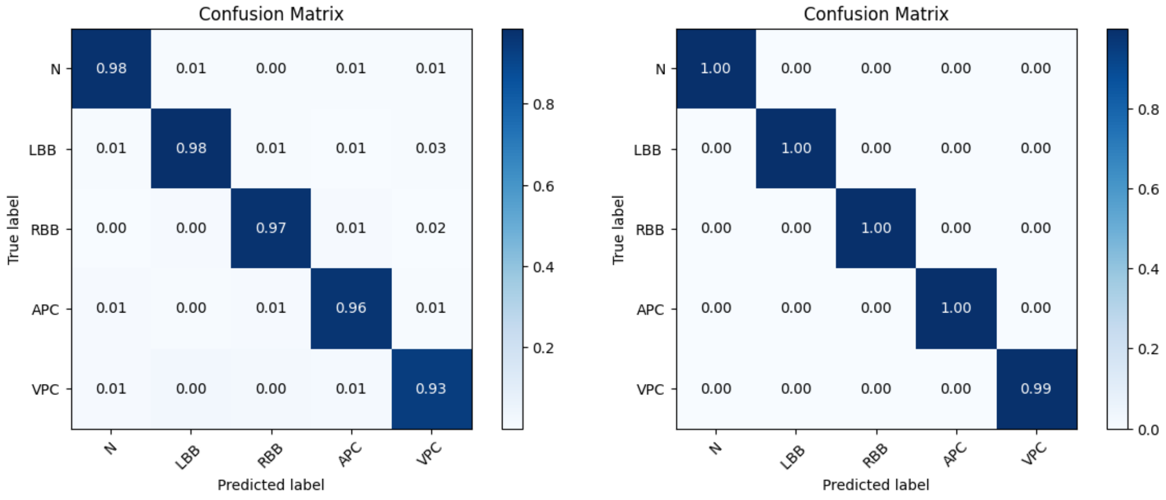
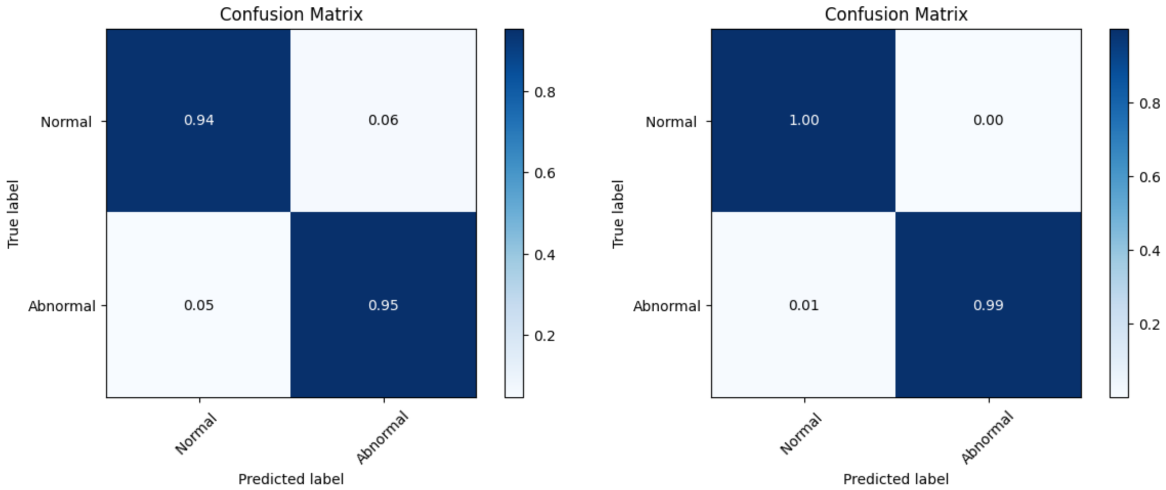


Fig. 6. Features visualization of the last layer



(a) 2D-CNN **(b) 2D-CNN + SAM**

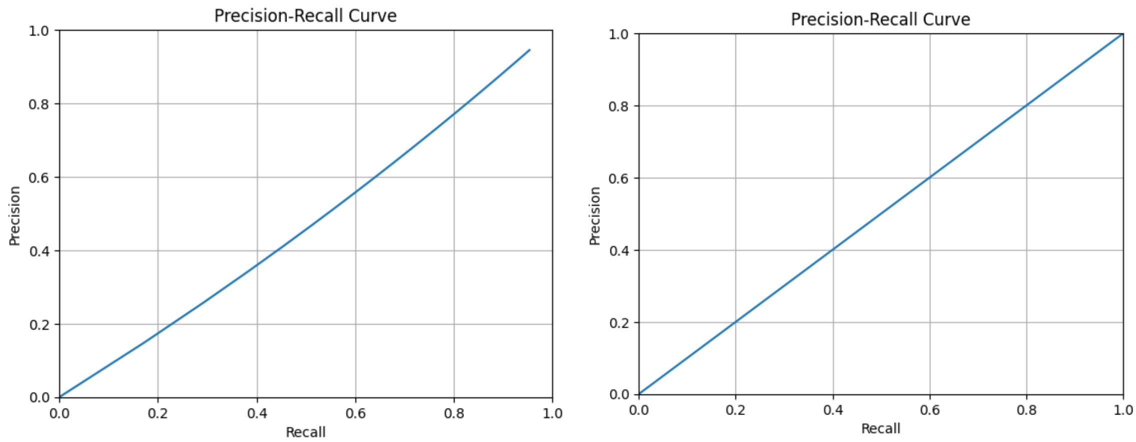
Fig. 7. Confusion matrices (multi-class)



(a) 2D-CNN **(b) 2D-CNN + SAM**

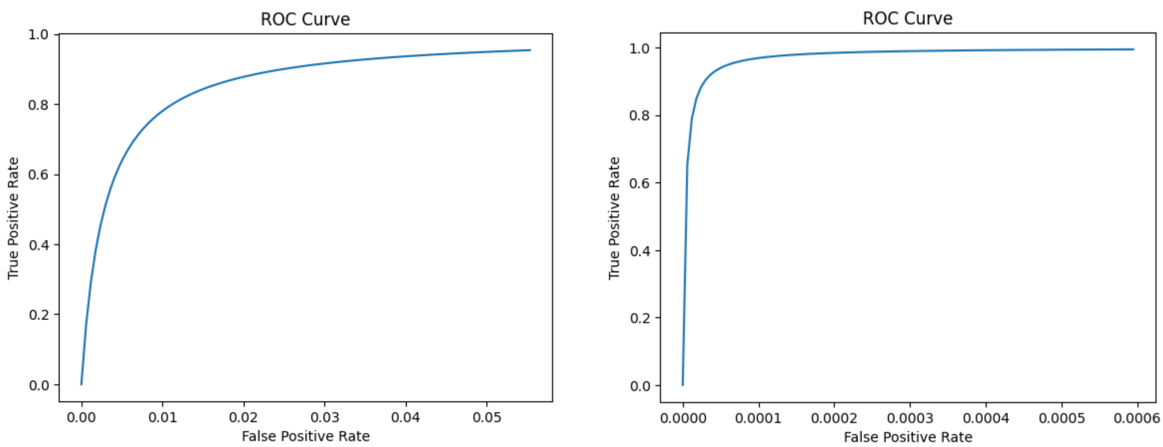
Fig. 8. Confusion matrices (Binary)

Precision-Recall and ROC curves are vital tools for evaluating classification model performance. The Precision-Recall curve shows how much precision and recall change concerning different decision thresholds, as shown in Fig. 9. On the other hand, the ROC is a plot of the trade-offs in true positive rates against the false positive rate at different thresholds, as shown in Fig. 10. Both the curves in isolation provide insights on a model's potential for discriminating between classes and help in choosing a proper threshold for classification tasks.



(a) 2D-CNN (b) 2D-CNN + SAM

Fig. 9. The Precision-Recall curves



(a) 2D-CNN **(b) 2D-CNN+SAM**

Fig. 10. The ROC curves

Table 4 summarizes the most relevant research efforts reviewed that have been carried out with the same database. Such collation helps provide input into the commonality of the individual works that have tremendously evolved the landscape of arrhythmia diagnosis using deep learning techniques with ECG signal analysis. These findings synthesize to illuminate the way ahead for further work in cardiac health monitoring and diagnostic protocols.

Table 4. Comparison with other related works

Years	Model	Class	Accuracy	Specificity	Sensitivity	Precision	F1 score
(Huang et al., 2019)	CNN	5	99%	-	-	-	-
(Asgharzadeh et al., 2020)	CNN	5	98.33%	99.09%	98.33%	98.34%	-
(Pandey et al., 2020)	LSTM	5	99.37%	99.14%	94.89%	96.73%	95.77%
(Olanrewaju et al., 2021)	CNN	3	98.7%	-	-	-	-
(Degirmenci et al., 2022)	CNN	5	99.70%	99.22%	99.70%	-	-
(Yared et al., 2023)	CNN	3	99.2%	99.6%	99.2%	-	-
2024	**This Work**	5	99.784%	99.944%	-	99.782%	99.776%
	CNN+SAM	2	99.70%	99.94%	99.47%	99.94%	99.70%

5. CONCLUSIONS

This study introduced an advanced methodology for diagnosing arrhythmias using ECG signals, employing a 2D CNN model enriched with Continuous Wavelet Transform (CWT) and a Selective Attention Mechanism (SAM). The SAM enhanced the model's ability to focus on crucial features in the ECG signals, contributing to improved classification accuracy. The model efficiently categorized ECG signals into Normal and Abnormal classes, with further discrimination among four specific subtypes within the Abnormal class. This comprehensive framework aimed to provide an accurate and detailed classification of diverse arrhythmia patterns in ECG data by integrating deep learning and signal processing techniques. The 2D CNN model with SAM achieved exceptional accuracies of 99.784% for multi-class classification and 99.94% for binary classification.

6. REFERENCES

Abdulaal, A. H, M. Valizadeh, Yassin, R. A., Albaker, B. M., Abdulwahhab, A. H., Amirani, M. C, & Shah, S. (2024d). Unsupervised histopathological sub-image analysis for breast cancer diagnosis using variational autoencoders, clustering, and supervised learning. Journal of Engineering and Sustainable Development, 28(6), 729–744.

Abdulaal, A. H., Yassin, R. A., Valizadeh, M., Abdulwahhab, A. H., Jasim, A. M., Mohammed, A. J., Jabir, H. J., Albaker, B. M., Dheyaa, N. H., & Amirani, M. C. (2024e). Cutting-Edge

CNN approaches for breast histopathological classification: The impact of spatial attention mechanisms. ShodhAI: Journal of Artificial Intelligence, 1(1), 109–130.

Abdulaal, A.H., Dheyaa, N.H., Abdulwahhab, A.H., Yassin, R.A., Valizadeh, M., Albaker, B.M. and Mustaf, A.S. (2024a). Deep Learning-based Signal Identification in Wireless Communication Systems: a Comparative Analysis on 3G, LTE, and 5G Standards. Al-Iraqia Journal for Scientific Engineering Research , 3(3), pp.66–70.

Abdulaal, A.H., Valizadeh, M., Albaker, B.M., Yassin, R.A., Amirani, M.C. and Shah, A.F.M.S. (2024b). Enhancing Breast Cancer Classification Using a Modified GoogLeNet Architecture with Attention Mechanism. Al-Iraqia Journal of Scientific Engineering Research, 3(1).

Abdulaal, A.H., Valizadeh, M., Amirani, M.C. and Shahen Shah, A.F.M. (2024c). A self-learning Deep Neural Network for Classification of Breast Histopathological Images. Biomedical Signal Processing and Control, 87(Part B), p.105418.

Abdulwahhab, A.H., Alaa Hussein Abdulaal, Assad, Ali Abdulwahhab Mohammed and Morteza Valizadeh (2024). Detection of Epileptic Seizure Using EEG Signals Analysis Based on Deep Learning Techniques. Chaos, Solitons & fractals/Chaos, Solitons and Fractals, 181, pp.114700–114700.

Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H. and Adam, M. (2017). Application of Deep Convolutional Neural Network for Automated Detection of Myocardial Infarction Using ECG Signals. Information Sciences, 415-416, pp.190–198.

Ahmed, A.A., Ali, W., Abdullah, T.A.A. and Malebary, S.J. (2023). Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model. Mathematics, 11(3), p.562.

Ardeti, V.A., Kolluru, V.R., Varghese, G.T. and Patjoshi, R.K. (2023). An Overview on state-of-the-art Electrocardiogram Signal Processing methods: Traditional to AI-based Approaches. Expert Systems with Applications, 217, p.119561.

Asgharzadeh-Bonab, A., Amirani, M.C. and Mehri, A. (2020). Spectral Entropy and Deep Convolutional Neural Network for ECG Beat Classification. Biocybernetics and Biomedical Engineering, 40(2), pp.691–700.

Cheikhrouhou, O., Mahmud, R., Zouari, R., Ibrahim, M., Zaguia, A. and Gia, T.N. (2021). One-Dimensional CNN Approach for ECG Arrhythmia Analysis in Fog-Cloud Environments. IEEE Access, 9, pp.103513–103523.

Chen, C., Hua, Z., Zhang, R., Liu, G. and Wen, W. (2020). Automated Arrhythmia Classification Based on a Combination Network of CNN and LSTM. Biomedical Signal Processing and Control, 57, p.101819.

Degirmenci, M., Ozdemir, M.A., Izci, E. and Akan, A. (2021). Arrhythmic Heartbeat Classification Using 2D Convolutional Neural Networks. IRBM, 43(5), pp.422–433.

Gaidai, O., Cao, Y. and Loginov, S. (2023). Global cardiovascular diseases death rate prediction. Current Problems in Cardiology, 48(5), p.101622.

Giancaterino, S., Urey, M.A., Darden, D. and Hsu, J.C. (2020). Management of Arrhythmias in Cardiac Amyloidosis. JACC: Clinical Electrophysiology, 6(4), pp.351–361.

Gutiérrez-Gnecchi, J.A., Morfin-Magaña, R., Lorias-Espinoza, D., Tellez-Anguiano, A. del C., Reyes-Archundia, E., Méndez-Patiño, A. and Castañeda-Miranda, R. (2017). DSP-based Arrhythmia Classification Using Wavelet Transform and Probabilistic Neural Network. Biomedical Signal Processing and Control, 32, pp.44–56.

Hao, W., Yizhou, W., Yaqin, L. and Zhili, S. (2020). The Role of Activation Function in CNN. 2020 2nd International Conference on Information Technology and Computer Application (ITCA).

Huang, J., Chen, B., Yao, B. and He, W. (2019). ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network. IEEE Access, 7, pp.92871–92880.

Janssens, A.C.J.W. and Martens, F.K. (2020). Reflection on Modern methods: Revisiting the Area under the ROC Curve. International Journal of Epidemiology, 49(4), pp.1397–1403.

Li, D., Zhang, J., Zhang, Q. and Wei, X. (2017). Classification of ECG Signals Based on 1D Convolution Neural Network. 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom).

Liu, Z., Chen, Y., Zhang, Y., Ran, S., Cheng, C. and Yang, G. (2023). Diagnosis of Arrhythmias with Few Abnormal ECG Samples Using metric-based Meta Learning. Computers in Biology and Medicine, 153, pp.106465–106465.

Martis, R.J., Acharya, U.R., Mandana, K.M., Ray, A.K. and Chakraborty, C. (2013). Cardiac Decision Making Using Higher Order Spectra. Biomedical Signal Processing and Control, 8(2), pp.193–203.

Mathur, P., Srivastava, S., Xu, X. and Mehta, J.L. (2020). Artificial Intelligence, Machine Learning, and Cardiovascular Disease. Clinical Medicine Insights: Cardiology, 14, p.117954682092740.

Mohammed, H. A., Kareem S. W., & Mohammed A. S. (2022). A comparative evaluation of deep learning methods in digital image classification. Kufa Journal of Engineering, 13(4), 53–69.

Mohonta, S.C., Motin, M.A. and Kumar, D.K. (2022). Electrocardiogram Based Arrhythmia Classification Using Wavelet Transform with Deep Learning Model. Sensing and Bio-Sensing Research, 37, p.100502.

Moody, G.B. and Mark, R.G. (2001). The Impact of the MIT-BIH Arrhythmia Database. IEEE Engineering in Medicine and Biology Magazine, 20(3), pp.45–50.

Myrovali E., Tachmatzidis D, Antoniadis, A., & Vassilikos V. (2022). Identifying patients with paroxysmal atrial fibrillation from sinus rhythm ECG using random forests. Expert Systems with Applications, 213(Part A), 118948–118948.

Olanrewaju, R.F., Ibrahim, S.N., Asnawi, A.L. and Altaf, H. (2021). Classification of ECG Signals for Detection of Arrhythmia and Congestive Heart Failure Based on Continuous Wavelet Transform and Deep Neural Networks. Indonesian Journal of Electrical Engineering and Computer Science, 22(3), p.1520.

Pandey, S.K. and Janghel, R.R. (2020). Automatic Arrhythmia Recognition from Electrocardiogram Signals Using Different Feature Methods with Long short-term Memory Network Model. Signal, Image and Video Processing, 14.

Talib, M., Abdullah, A. A., K. Abdullah, A., & Bahaa, B. (2020). Twin fetus ECG signal extraction based on temporal predictability. Kufa Journal of Engineering, 11(1), 35–51.

Ullah, A., Anwar, S.M., Bilal, M. and Mehmood, R.M. (2020). Classification of Arrhythmia by Using Deep Learning with 2-D ECG Spectral Image Representation. Remote Sensing, 12(10), p.1685.

Ullah, A., Rehman, S. ur, Tu, S., Mehmood, R.M., Fawad and Ehatisham-ul-haq, M. (2021a). A Hybrid Deep CNN Model for Abnormal Arrhythmia Detection Based on Cardiac ECG Signal. Sensors, 21(3), p.951.

Ullah, W., Siddique, I., Zulqarnain, R.M., Alam, M.M., Ahmad, I. and Raza, U.A. (2021b). Classification of Arrhythmia in Heartbeat Detection Using Deep Learning. Computational Intelligence and Neuroscience, 2021, pp.1–13.

Wang, Y., Yang, G., Li, S., Li, Y., He, L. and Liu, D. (2023). Arrhythmia Classification Algorithm Based on multi-head self-attention Mechanism. Biomedical Signal Processing and Control, 79, pp.104206–104206.

Yared Daniel Daydulo, Bheema Lingaiah Thamineni and Ahmed Ali Dawud (2023). Cardiac Arrhythmia Detection Using Deep Learning Approach and Time Frequency Representation of ECG Signals. BMC Medical Informatics and Decision Making, 23(1).

Yıldırım, Ö., Pławiak, P., Tan, R.-S. and Acharya, U.R. (2018). Arrhythmia Detection Using Deep Convolutional Neural Network with Long Duration ECG Signals. Computers in Biology and Medicine, 102, pp.411–420.