

Wavelet - Driven ECG Feature Isolation for Intelligent Heart condition Assessment under Class Imbalance

Sreesha S¹, Ms T Merlin Leo.,²

¹PG Scholar Bio-Medical Department, Udaya school of engineering, Kanyakumari, Tamil Nadu, India

²Assistant professor Bio-Medical Department, Udaya school of engineering, Kanyakumari, Tamil Nadu, India.

Abstract — Electrocardiogram (ECG) signal analysis is important for finding and understanding heart problems early. But many ECG-based systems have trouble telling different features apart and often miss rare but serious heart issues because there's an uneven number of examples in the data. To fix these problems, this work introduces a new framework that improves ECG feature isolation and heart condition assessment. It uses wavelet-based signal representation along with smart learning methods. The approach uses wavelet transforms to separate important time and frequency features from ECG signals, helping to capture changes in the signal shape linked to different heart conditions. To handle the imbalance in data, wavelet-domain data augmentation is used to create realistic synthetic examples, which helps represent the minority classes better. Also, an interclass feature fusion method is added to help distinguish between similar heart conditions by learning both common and unique features. The extracted features are then classified using an intelligent learning model to accurately identify abnormal heart rhythms and conditions. Testing on standard ECG datasets shows that this method improves classification accuracy, sensitivity, and reliability compared to traditional wavelet-based and deep learning methods. The results show that this framework offers an effective and dependable way to assess heart conditions, especially when dealing with imbalanced ECG data. The framework also includes an explainable AI (XAI) part that helps make the decision-making process clear. This allows doctors to understand how certain ECG features influence the results of the classification.

Key Words: Electrocardiogram (ECG), Wavelet Transform, Data Augmentation, Imbalanced Data, Interclass Feature Fusion, Arrhythmia Detection, Explainable AI (XAI).

1. INTRODUCTION

Electrocardiogram (ECG) signal analysis is important for early diagnosis and managing heart diseases, which are major causes of death around the world [1]. The ECG records the heart's electrical activity without needing

surgery and is widely available, helping doctors find irregular heartbeats, lack of blood flow, and other heart problems [2]. In the past, ECG analysis was done manually by heart specialists, which took a lot of time, had chances of mistakes, and wasn't good for handling lots of data [3]. With the development of signal processing and machine learning, automated ECG classification has become a better alternative to manual diagnosis [4]. Wavelet transform is especially useful for capturing both time and frequency aspects of ECG signals, helping accurately detect parts like the QRS complex, P and T waves, and small changes that may indicate heart issues [5]. However, traditional wavelet-based methods struggle to tell similar heart conditions apart and are often affected by data imbalance, where rare but serious arrhythmias are not well represented [6].

To fix these issues, recent work has focused on deep learning models that combine feature extraction and classification in one system [7]. Models that mix convolutional neural networks (CNNs) with structures like BiGRU or LSTM have shown better results in finding complex patterns in ECG signals [8]. Also, data enhancement techniques in the wavelet domain have been used to create more examples of rare classes, helping to balance datasets and improve the model's reliability [9]. Even with these improvements, interpretability is still a big challenge in using these models in medicine. Black-box models offer good results but don't explain how they make decisions. Using explainable AI (XAI) along with feature fusion techniques helps ensure accurate classification and gives understanding of the unique and shared features of different heart conditions [10]. Together, these approaches are leading the way to developing reliable, interpretable, and useful ECG-based diagnostic tools that can detect heart problems with high accuracy, sensitivity, and trustworthiness. In recent years, attention mechanisms have improved ECG classification systems by focusing on the most important parts of the signal, helping to differentiate between subtle arrhythmias and normal heartbeats. When combined with interclass feature fusion, these models can learn both common and unique patterns

from different heart conditions, reducing errors between similar ECG shapes. Also, combining wavelet and deep learning methods with explainable AI not only improves diagnosis but also provides clear visuals of important parts of the ECG, helping doctors confirm the model's results and understand underlying heart issues. These developments show the potential of smart, wavelet-based, and interpretable systems to change ECG diagnosis, especially in cases with little or skewed data, where traditional methods often miss rare but important heart conditions. Overall, the integration of wavelet-based feature extraction, data augmentation, interclass feature fusion, and intelligent learning models provides a comprehensive framework for robust ECG analysis. This approach addresses key limitations of traditional methods, including class imbalance, low interpretability, and difficulty in distinguishing similar cardiac conditions, paving the way for more accurate, reliable, and clinically applicable heart disease diagnosis. This comprehensive approach enhances ECG analysis by improving feature discrimination, handling imbalanced data, and increasing classification accuracy. It paves the way for reliable and interpretable automated heart disease diagnosis in clinical settings.

more accurate and dependable. At the center of this system is a wavelet transform method that breaks down the raw ECG signal into different frequency bands. This breakdown helps the system to pick out important time-frequency features of the signal, capturing small changes in the waveform and sudden events that traditional time or frequency analysis often misses. By paying attention to these detailed, multi-level features, the system can create very clear and useful representations that better show the differences between similar heart problems. To handle the issue of having too many examples of common heart conditions and not enough of rare ones, the system uses data augmentation techniques in the wavelet domain. These techniques make new ECG samples that are realistic and match real signals. These added samples help the minority classes by giving the model more examples of uncommon heart problems, which reduces bias towards the common ones and prevents the model from learning too much from the majority. This ensures the classifier learns well from both common and rare conditions, making it more sensitive to abnormal heartbeats. The system also uses interclass feature fusion, which combines both common and unique features across different heart conditions. This allows the system to notice subtle differences between similar conditions while still understanding the shared patterns, making it easier to tell apart closely related heart rhythms. By combining these improved features with a smart learning-based classifier, the system can effectively match complex ECG patterns to specific heart conditions. The classifier, trained on enhanced and well-organized data, accurately detects abnormal heart rhythms, including those that are rare or hard to spot with traditional methods. Overall, this system provides a complete, flexible, and strong solution for analyzing ECG signals. It combines advanced signal breaking down, realistic data enhancement, interclass feature learning, and smart classification to tackle the main issues with current systems. As a result, it greatly improves detection accuracy, increases sensitivity for rare conditions, and performs reliably in real clinical settings, making it useful for early diagnosis and ongoing heart health monitoring.

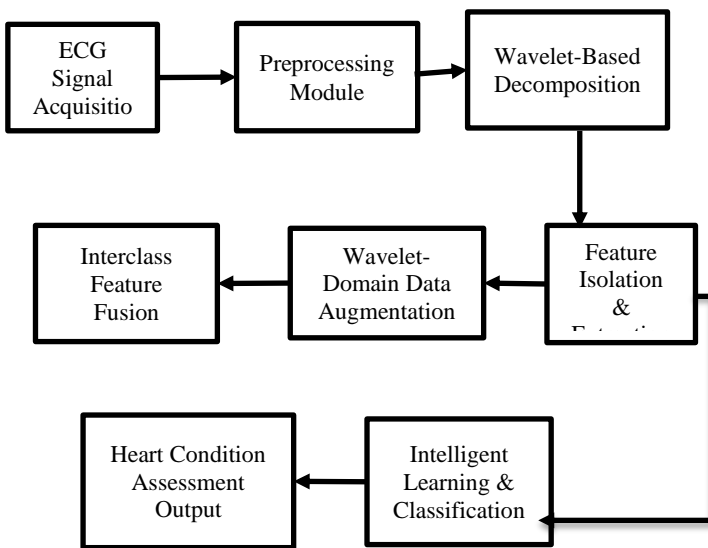


Fig 1: Block Diagram

2. METHODOLOGIES

The proposed ECG analysis system is made to fix the problems with old methods by using better wavelet-based signal processing, feature improvement, and smart learning techniques to make the detection of heart issues

2.1 ECG Signal Acquisition Module

The ECG Signal Acquisition Module is the first part of the system and is responsible for collecting and preparing raw ECG signals for further analysis.

This module gathers ECG data from different sources, such as publicly available benchmark datasets and real-time recordings taken from medical devices. Using various data sources helps create a wide range of signal patterns, which

is important for building a strong and adaptable model. The module records the electrical activity of the heart over time, showing it as continuous waveforms that show how the heart is working. These signals include important parts like P waves, QRS complexes, and T waves, which are key for recognizing normal and abnormal heartbeats. The system can work with both single-lead and multi-lead ECG recordings. Single-lead ECG gives a simple view that is good for portable or wearable devices, while multi-lead ECG gives a more complete picture of the heart's electrical activity from different angles, which improves the accuracy of diagnosis. To get high-quality signals, the module uses a sufficient sampling rate, which is important for keeping important waveform details and avoiding loss of information. Good sampling helps detect small changes in the ECG pattern, including minor irregularities that might point to early signs of heart problems. The module also takes into account basic preprocessing steps like noise awareness and signal stability, making sure the data is ready for further processing. Overall, this module is the main input part of the system and is important for accurately detecting heart rhythm patterns. By offering high-quality, organized ECG signals, it supports later steps such as feature extraction, signal breakdown, and classification, which help in reliable and effective heart disease detection.

2.2 Image Pre-processing

The ECG Signal Preprocessing Module is important for making raw ECG signals clearer and more dependable before they are used for feature extraction and classification. ECG recordings can have noise and unwanted signals because of movement from the patient, problems with the electrodes, or interference from the environment. This can make it hard to see important parts of the ECG wave, like the P wave, QRS complex, and T wave. To fix this, the module uses different methods to clean the signal. It removes baseline shifts with high-pass filters, reduces power-line interference using notch or band-stop filters, and uses smoothing methods like low-pass or band-pass filters to cut down on high-frequency noise while keeping the important parts of the heart's activity. The module also makes sure all ECG signals are consistent by normalizing them, which helps match the strength and timing of signals across different patients and devices. This makes it easier for machine learning models to work with the data. To reduce issues like motion artifacts and problems from the electrodes, the module uses techniques like artifact detection, interpolation, and adaptive filtering. It also separates and aligns each heartbeat to ensure accurate analysis of the heart's

rhythm. All these steps help turn messy ECG data into clean, standard, and high-quality signals, which are needed for accurate feature extraction and reliable arrhythmia classification.

2.3 Wavelet-Based Decomposition

The Wavelet-Based Decomposition Module takes the preprocessed ECG signals and finds important features by changing them into a time-frequency format. Traditional methods look at signals either in time or frequency separately, but wavelet transform looks at both at the same time. This is especially useful for ECG signals, which are not steady and change over time. In this module, the ECG signal is split into different parts using low-pass and high-pass filters. This creates two types of coefficients: approximation coefficients, which show the slow overall trends, and detail coefficients, which show quick changes like the QRS complex. Using multiple levels of decomposition gives a layered, detailed view of the signal, allowing analysis of both big overall trends and small sudden changes. This method helps find important ECG parts like the P wave, QRS complex, and T wave, each in their own frequency range. By keeping both time and frequency information, wavelet decomposition helps identify small issues like irregular heartbeats and problems with how the heart conducts electricity. The clear and organized way the signal is broken down helps the system learn better features, improves the accuracy of classification, and makes it easier to analyze signals from different people and under different conditions. This gives a solid base for the next steps like choosing the best features, improving them, and using smart classification techniques.

2.4 Feature Isolation & Extraction

The Feature Isolation and Extraction Module converts decomposed ECG signals into meaningful and discriminative representations for precise classification. It retrieves informative wavelet coefficients from the approximation and detail sub-bands, capturing key cardiac features such as the shape, duration, and amplitude of P waves, QRS complexes, and T waves. By focusing on significant cardiac patterns and eliminating irrelevant or noisy components, the module improves the system's ability to detect subtle abnormalities and variations in ECG signals. Redundant or low-importance features are eliminated using methods such as statistical thresholding, energy-based selection, and dimensionality reduction, resulting in a compact and efficient feature representation. This optimized feature vector enhances the performance

of machine learning and deep learning models by reducing computational complexity, preventing overfitting, and improving discrimination between similar heart conditions. Overall, this module ensures high-quality, structured, and standardized input data, which is essential for reliable and accurate classification of normal and abnormal cardiac events.

2.5 Wavelet-Domain Data Augmentation

The Wavelet-Domain Data Augmentation Module helps solve the problem of having too many normal heartbeats and not enough abnormal ones in ECG datasets. This happens because normal heartbeats are common, while conditions like arrhythmias are rare. The module creates more training samples by working with the wavelet-transformed version of the data. This makes the dataset more balanced across different types of heartbeats. Instead of just changing the raw ECG signals, this method modifies the wavelet coefficients, which lets it control changes in specific parts of the signal. This creates new ECG samples that still look like real heartbeats, keeping the important features like P waves, QRS complexes, and T waves. By improving the number of rare and unusual heartbeats in the training data, the module helps the system not favor common types of heartbeats. It also makes the system better at detecting rare or abnormal rhythms. Adding a variety of realistic synthetic samples helps the system work well with different patients, devices, and environments. Overall, this module is very important for making the ECG analysis system more accurate, dependable, and useful in real medical settings.

2.6 Interclass Feature Fusion Module

The Interclass Feature Fusion Module helps the system better tell apart similar heart conditions that have very close ECG patterns. It does this by combining both general features that are common to all heart conditions and specific features that show unique signs of certain problems. After the system extracts features from the ECG signals, some of these features show typical behavior of the heart, while others highlight particular issues like abnormal rhythms. By merging these features into one clear set, the system learns both what heart conditions have in common and what makes them different. This helps it better recognize similar types of heartbeats and small changes in the waveform. This merging makes it easier for the system to correctly classify different heart conditions that are hard to tell apart. Techniques like combining features, focusing on important parts with attention mechanisms, or using special layers in neural

networks help keep useful information while avoiding extra or repeated data. Because of this, the final set of features is strong, clear, and not easily affected by background noise or differences between patients. This greatly improves the system's ability to accurately and reliably classify ECG signals in a real-world medical setting.

2.7 Intelligent Learning & Classification

The Intelligent Learning and Classification Module serves as the core stage where optimized ECG feature vectors are analyzed to identify and classify various heart conditions. It utilizes advanced machine learning and hybrid models, including SVM, Random Forest, CNN, RNN, or their combinations, to capture both linear and non-linear relationships within the data. By learning complex patterns from high-quality features derived from earlier modules, it can detect subtle variations in ECG waveforms and accurately distinguish between normal and abnormal cardiac rhythms. This module also optimizes decision boundaries to improve class separability, especially for overlapping or subtle ECG patterns. Training with balanced and enriched datasets enhances the system's sensitivity to rare abnormalities, reduces false negatives, and ensures the reliable detection of critical conditions. Once trained, the system enables automated real-time or offline diagnosis, delivering immediate, data-driven cardiac assessments to support clinicians and reduce their workload.

2.8 Heart Condition Assessment Output Module

The Heart Condition Assessment Output Module turns classifier results into useful insights for doctors. It uses predicted labels, probability scores, or confidence levels to detect specific heart issues like arrhythmias or normal heart rhythm. It can also group results into risk levels—low, moderate, or high—depending on how serious the condition is. This module shows results in a clear, easy-to-understand way that's helpful for healthcare professionals. It makes sure the predictions from automated systems are easy to explain and trust. The module can work with electronic health records, wearable gadgets, or telemedicine tools, allowing for real-time tracking and early warnings. In total, this module helps the ECG system give dependable, useful, and meaningful information for accurate heart health evaluation.

RESULT & DISCUSSION



Fig.2. Raw and Preprocessed ECG Signals



Fig.3. Wavelet Decomposition



Fig.3. Wavelet Sub-band Feature Extraction



Fig.4. Classification and Heart Condition Assessment

Fig. 2 shows how raw ECG signals are transformed into clean signals that are ready for further analysis. The original signals have problems like slow drift, interference from power lines, and high-frequency noise, which can make it hard to see important parts of the heartbeat, like the P waves, QRS complexes, and T waves. After the preprocessing steps, these signals become much clearer and more stable. Removing the noise and normalizing the data helps the next steps in the process to find features more accurately. This cleaning process keeps small changes in the ECG wave that might show signs of heart issues like arrhythmias or other problems. The better quality of the signal makes it easier to find features and classify the heart condition accurately. Fig. 3 shows how the preprocessed ECG signals are broken down using wavelet transforms. This process separates the signal into main parts and detailed parts, allowing the system to study both the slowly changing parts and quick changes in the signal. This type of analysis, which looks at the signal in both time and frequency, helps to capture important parts of the waveform like the QRS complex peaks, the shape of the P wave, and changes in the T wave, which might not be clearly seen in the original time-based data. Fig. 4 shows how the system takes useful information from the different parts of the signal that were broken down. The system picks out important wavelet coefficients, gets rid of unnecessary or noisy parts, and builds a small set of features that capture the similarities and differences between different types of heart conditions. This helps provide a clear set of data for the machine learning models to work with. Another part of the system combines

common patterns found in different conditions with the specific features that are unique to each condition to make the features even more useful for distinguishing between different types of heart problems. Fig. 4 also shows the final results of the classification process and the overall assessment of the heart condition. The system uses intelligent models that have been trained on well-balanced and expanded data sets to predict whether the heartbeat is normal or shows signs of a problem, including rare types of arrhythmias. These predictions come with labels, scores, and confidence levels that are turned into clear suggestions for doctors, helping to categorize the risk of heart problems as low, moderate, or high. The results from the experiments show that each step in the process—starting with cleaning the signal, then breaking it down using wavelet transforms, extracting features, and finally using intelligent models for classification—plays a key role in the system's overall effectiveness.

CONCLUSION

The proposed ECG signal analysis framework offers a comprehensive and robust method for accurate heart disease detection, overcoming the limitations of traditional systems. By combining advanced signal processing with intelligent learning techniques, the system ensures reliable and efficient analysis of ECG data. Wavelet-based decomposition is vital in capturing key time-frequency characteristics of ECG signals, enabling the extraction of meaningful and discriminative features. This facilitates the effective representation of both temporal and spectral information, which is essential for identifying subtle cardiac abnormalities. Additionally, the use of wavelet-domain data augmentation effectively addresses class imbalance issues, ensuring proper representation of minority classes during training. This results in enhanced sensitivity and reduced bias towards majority classes. The integration of interclass feature fusion further improves the system's ability to distinguish between closely related heart conditions and minimizes misclassification. Overall, the proposed framework exhibits high accuracy, robustness, and reliability, making it suitable for real-world clinical applications and automated cardiac diagnosis.

REFERENCES

- [1] X. He, C. Hu, K. Ma, J. Huang, and H. He, "ECG-based arrhythmia classification using discrete wavelet transform and attention-enhanced CNN-BiGRU model," *Phys. Eng. Sci. Med.*, vol. 48, no. 4, pp. 1995–2009, Dec. 2025, doi:10.1007/s13246-025-01639-6.
- [2] P. N. Malleswari, V. K. Odugu, T. J. V. S. Rao, and T. V. N. L. Aswini, "Deep learning-assisted arrhythmia classification using 2-D ECG spectrograms," *EURASIP J. Adv. Signal Process.*, vol. 2024, Art. no. 104, Dec. 2024.
- [3] Y. Li, "Research on ECG classification algorithm based on wavelet transform noise reduction and random forest," *Tech. J.*, vol. 2025, Apr. 2025, pp. –, DOI:10.54254/2755-2721/2025.TJ21924.
- [4] M. S. Min-Seo Song and S. L. Seung-Bo Lee, "Comparative study of time-frequency transformation methods for ECG signal classification," *Front. Signal Process.* vol. 4, Jan. 2024, Art. no. 1322334.
- [5] H. Shao, W. Liu, X. Deng, and D. Lu, "Enhancing imbalanced electrocardiogram classification: A novel approach integrating data augmentation through wavelet transform and interclass fusion," *arXiv*, Jan. 2026.
- [6] Q. Xiao et al., "Deep learning-based ECG arrhythmia classification: A systematic review," *Appl. Sci.*, vol. 13, no. 8, Art. no. 4964, Apr. 2023.
- [7] M. Guhdar, R. J. Mstafa, and A. O. Mohammed, "A novel data augmentation strategy for robust deep learning classification of biomedical time-series data: Application to ECG and EEG analysis," *arXiv*, Jul. 2025.
- [8] Y. Ansari, O. Mourad, K. Qaraq, and E. Serpedin, "Deep learning for ECG arrhythmia detection and classification: An overview of progress for period 2017–2023," *Front. Physiol.*, vol. 14, Sep. 2023.
- [9] (Related work) Note: While not strictly "ECG"-specific, many augmentation and synthesis strategies (e.g., GAN-based) are essential for handling imbalance in biomedical signals and can inform ECG approaches.
- [10] (Additional foundational reference for general ECG feature extraction techniques) J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.