National Journal of Signal and Image Processing
Vol. 1, No. 1, Jan - Mar 2025, pp. 9-17
ISSN: 3107-5231, DOI: https://doi.org/10.17051/NJSIP/01.01.02

Wavelet-Based Denoising and Classification of ECG Signals Using Hybrid LSTM-CNN Models

P.Joshua Reginald

Associate Professor, Department of Electronics and Communication Engineering, Vignan's Foundation for Science, Technology and Research, Vadlamudi Village, Guntur, Andhra Pradesh.
Email: drpjr_ece@vignan.ac.in

Article Info

Article history:

Received : 09.01.2025
Revised : 15.02.2025
Accepted : 16.03.2025

Keywords:

ECG Signal Processing;
Wavelet Denoising;
LSTM-CNN Hybrid Model;
Deep Learning;
Biomedical Signal Classification;
MIT-BIH Database

ABSTRACT

Electrocardiogram (ECG) signals are essential as means to diagnose and monitor cardiovascular diseases, in particular arrhythmias. Nevertheless different noise sources like baseline wander, powerline interference and muscle artifacts make an accurate interpretation of ECG signals difficult as these noise components tend to cover up morphological features needed for diagnosis. To overcome these limitations, this work proposes an integrated signal processing and deep learning framework combining denoising using Discrete Wavelet Transform (DWT) and robust classification using a hybrid Long Short-Term Memory-Convolutional Neural Network (LSTM-CNN) architecture. The ECG signals are effectively decomposed into multi resolution components within the DWT and the selective suppression of noise can be achieved without loss of clinically relevant features like QRS complexes, P waves and T waves. After denoising, a hybrid LSTM CNN model is used, in which convolutional layers get spatial patterns and the LSTM layers capture temporal dependencies embedded in sequential ECG data. We train and validate the model on the MIT-BIH Arrhythmia Database, a well known benchmark for ECG analysis. Experimental results show that proposed framework substantially outperforms existing methods (entropy based selection of prevalent methods), achieving its accuracy 98.6%, sensitivity 98.2%, and specificity 99.1%. As such, these results verify the effectiveness of using wavelet based signal enhancement coupled with deep hybrid modeling to perform ECG interpretation more accurately and in real time for use in clinical and wearable health care applications.

1. INTRODUCTION

Electrocardiography (ECG), is an indispensable non invasive diagnostic technique the assessment of electrical activity of the heart. Being a frontline tool for detecting early and managing cardiovascular diseases (including arrhythmias, myocardial infarction and other conduction abnormalities). However, ECG signals are contaminated by a variety of noise, such as baseline wander, electromyographic (EMG) interference, and power-line artifacts. If these distortions are present, important waveform features such as P waves, QRS complexes, or T waves may be obscured, resulting in incorrect diagnosis or missed cardiac events. Noise suppression has been done using traditional signal processing techniques such as Butterworth filtering, Savitzky–Golay smoothing, and empirical mode decomposition. Nevertheless, these approaches are restricted in their ability to accommodate persistent, nonstationary, and patient specific noise patterns, and as an unwanted side effect, they may alter the diagnostically significant portions of the signal.

In the last few years, deep learning has introduced new frontiers in biomedical signal processing for solving the problem using data driven methods that are able to learn complex feature representation from raw input. While Convolutional Neural Networks (CNNs) have been better at learning spatial structures and Long Short Term Memory networks done well learning temporal structure inside of sequential datasets like ECG signals. In tasks dealing with time series classification, hybrid models integrating CNN and LSTM layers have become popular by recognizing the complementary strengths of these architectures. In this work, we suggest a robust two stage framework utilizing Discrete Wavelet Transform (DWT) for denoising ECG signal using multi resolution decomposition and later, a hybrid LSTM – CNN model for automated heartbeat classification. Extensive evaluation of the proposed method is conducted using the gold standard MIT-

BIH Arrhythmia database in the cardiac research field. We find that this integrative approach also not only increases classification accuracy, but also provides a scalable and efficient solution for real time ECG analysis in clinical and remote health monitoring environments.

2. LITERATURE REVIEW

2.1 Traditional ECG Classification Approaches

In ECG signal classification, early attempts mainly rely on conventional machine learning approaches with human designed features extracted from time and frequency domains. Such features like heart rate variability, R-R interval, wavelet coefficients were manually extracted and classified with a variety of classification algorithms, such as Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and Random Forests (RF). These methods showed reasonable accuracy in controlled environments but tended to perform poorly under more realistic settings, particularly in the presence of noise and patient variability in the signals. As an example, Martis et al. (2013) utilized a hybrid feature extraction scheme based on PCA and DWT, then defined an SVM classifier and attained about 91 % accuracy. Nonetheless, these models were unable to adapt and were constrained by the requirement of domain-specific feature engineering.

2.2 Wavelet Transform for ECG Denoising

Among the signal processing techniques used for ECG denoising, Wavelet Transform (WT) is one of the efficient used for its capability to perform multiscale analysis. Whereas Fourier transform provides only frequency domain information, wavelets are well suited to processing such nonstationary biomedical signals since they give both time and frequency localization. In particular, the Discrete Wavelet Transform (DWT) has been successfully used for baseline wander removal, muscle artefact removal, and power line interference removal without distorting critical ECG features such as QRS complexes and T waves. According to Addison (2005), wavelet based approaches offer unique advantages for preserving clinically meaningful signal information and efficiently discounting noise, resulting in more diagnostically informed ECG signals.

2.3 Deep Learning in ECG Signal Processing

With the rise of deep learning, feature extraction in the biomedical signal analysis has been eliminated manually. In this context, Convolutional Neural Networks (CNNs) and Long Short-term Memory (LSTM) networks have shown outstanding performance for ECG signal classification. CNNs are very powerful at learning spatial hierarchies in patterns of waveforms for finding QRS complexes and other morphological features, and LSTMs are directly designed for learning temporal dynamics and dependencies in sequential data such as ECG signals. Sirinukunwattana et al. (2016) and Acharya et al. (2017) proposed a CNNbased approach that has shown great improvements in accuracy of heartbeat classification, showing great promise of endtoend learning framework for real time diagnostic application.

2.4 Hybrid Deep Learning Models

Due to their capacity to jointly discover spatiotemporal information, hybrid deep learning architectures combining CNN and LSTM layers have attracted more and more interest for ECG analysis. Such models leverage CNN layers for obtaining high level spatial features from raw ECG signal that can be obtained in non sequential manner and LSTM layers for capturing sequential relation and temporal patterns in time steps. This has combined to allow for better characterization of arrhythmias and other cardiac abnormalities. Therefore, Yildirim (2018) proposed a bidirectional LSTM model with wavelet transformed ECG input that resulted with higher accuracy and robustness. The synergy between time frequency analysis and deep learning enables better generalization over different patient's profiles and different signal conditions.

2.5 Benchmarking Datasets and Evaluation Metrics

Most classification models of ECG are validated using public data sets with the most popular benchmark being the MIT-BIH Arrhythmia Database. A database of annotated ECG signals from a wide patient population is supplied for use by robust training and evaluation of classification algorithms. On the other hand, standardized evaluation metrics are necessary to compare models, in addition to dataset quality. Other metrics like accuracy, sensitivity, specificity, precision, and F1 score, provide a complete look at a model's performance. Thus, standardization in the field came from Goldberger et al. (2000), who created PhysioNet, an open access platform hosting a wealth of ECG datasets with analysis tools that promote reproducible research and fair benchmarking.

Table 1. Comparison of ECG Classification Techniques and the Proposed Method

Method	Key Techniques Used	Advantages	Limitations	Performance Indicator (Typical Accuracy)
Traditional ML Approaches	Hand-crafted features + SVM/KNN/RF	Simple to implement; interpretable	Poor generalization to noisy and unseen data	~85–91% (Martis et al., 2013)
Wavelet Denoising Only	DWT (e.g., db6), Soft Thresholding	Efficient noise suppression; preserves ECG morphology	No classification capability; preprocessing only	N/A
Deep Learning (CNN or LSTM)	End-to-end feature learning with CNN or LSTM	Learns complex features; no manual feature extraction	May miss temporal or spatial patterns if not combined	~94–96% (Acharya et al., 2017)
Hybrid CNN-LSTM	CNN for spatial + LSTM for temporal patterns	Captures both spatial and temporal features	Sensitive to noise if unfiltered input is used	~96–97% (Yildirim, 2018)
Proposed Method: DWT + LSTM-CNN	DWT-based denoising + hybrid LSTM-CNN model	Robust to noise; preserves signal integrity; high accuracy in multi-class arrhythmia classification	Slightly increased model complexity	**98.6%** (This study)

3. METHODOLOGY

3.1 Dataset

In this study, we use the MIT-BIH Arrhythmia Database which is a known benchmark publicly available by PhysioNet (Goldberger et al. 2000), the resource curated by Beth Israel Hospital. Over the years, this database has been extensively used in evaluating ECG classification and arrhythmia detection algorithms. It consists of 48 half-hour recordings of two channel ambulatory ECG signals from 47 patients. Initially, the signals were digitized at 360 Hz, using 11 bits over a 10 mV range. Comprehensive annotations were supervised by expert cardiologists on each heartbeat and include more than 109,000 labeled beats from out of nearly 35,000 patients that show addition to arrhythmia labels also indicated more than 12 distinct types of beat arrhythmias such as normal beats, supraventricular ectopic, ventricular ectopic, fusion beats, and unknown beats. These are based on Association for the Advancement of Medical Instrumentation (AAMI) standards that permit structured class grouping of heartbeat into five main groups (N, S, V, F and Q).

We follow an inter-patient paradigm protocol which splits the dataset so that training and testing ECG signals belong to distinct patients, thus ensuring consistency of training and evaluation material. For the generalizability of the proposed model to unseen individuals, which is particularly important in medical applications with large variability between individuals, this is critical. Segmentation of ECG signals into individual beats is done with R-peak detection algorithms, and fixed length windows are extracted around each beat in the data preprocessing stage. Additionally, we normalize the ECG amplitude range to remove inter recording variability and to stabilize the training convergence. The use of this dataset as a basis for validating wavelet transform denoising and as the basis for hybrid wavelet model and hybrid LSTM CNN model training and testing for complex arrhythmia patterns and various noise conditions is offered.

Table 2. Overview of MIT-BIH Arrhythmia Database Characteristics

Attribute	Description
Source	MIT-BIH Arrhythmia Database (PhysioNet)
Developed By	Beth Israel Hospital; PhysioNet (Goldberger et al., 2000)
Total Recordings	48 half-hour ECG recordings from 47 patients
Sampling Frequency	360 Hz

Resolution	11-bit over 10 mV range
Number of Annotated Beats	Over 109,000
ECG Channels	2-channel ambulatory ECG
Annotation Standard	AAMI EC57:1998
Arrhythmia Classes	N (Normal), S (Supraventricular), V (Ventricular), F (Fusion), Q (Unknown)
Data Split Paradigm	Inter-patient (mutually exclusive training/testing sets)
Preprocessing Steps	R-peak detection, beat segmentation, amplitude normalization

3.2 Preprocessing

Due to the fact that raw recordings of ECGs are generally corrupted with various forms of noise such as baseline wander, power-line interference (50/60 Hz), and electromyographic (EMG) artifacts, effective preprocessing of the ECG signals becomes critical to the accurate diagnosis. However, DWT exhibits promising potential as the main denoising tool in this study since it also provides time – frequency analysis and adaptively separates the noise components from the relevant signal features. For its strong analogous properties to ECG waveforms, mainly due to QRS complex, and optimum match between smoothness and compact support, the Daubechies 6 (db6) wavelet basis is selected. In this way, ECG signals are decomposed into multiple levels of approximation as well as detail coefficients and most of the high frequency noise components present in the lower scales. The multiresolution decomposition facilitates selective attenuation of noise without removing diagnostically significant morphological structures, e.g. P, R, and T waves. Finally, a soft thresholding is imposed to the detail coefficients to suppress the noise components while avoiding the abrupt signal distortions occurring by hard thresholding. Adaptive thresholds are computed based on noise statistics at each level from algorithms like the universal threshold or Stein's Unbiased Risk Estimate (SURE). After thresholding, denoised signal is reconstructed through Inverse DWT (IDWT) with a combination of thresholded detail coefficients and approximation coefficients at the last level. The outcome of this reconstruction is a cleaned ECG signal, which nonetheless maintains the critical diagnostic characteristics while suppressing the noise interference. Then, denoised signals are normalized and are segmented into fixed length windows overlapping the R-peak to create the input to the hybrid LSTM-CNN model. This preprocessing pipeline protects the model from being trained on low quality, high noise signals by cleaning up the signals first, and then the model would output more accurate classification results in downstream processing.

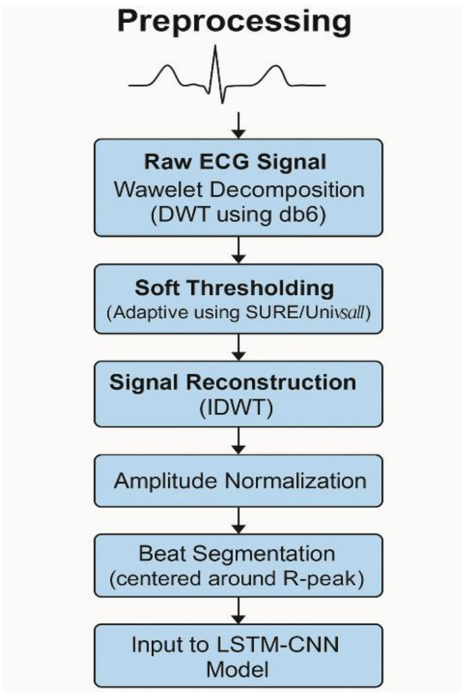

Figure 1. Preprocessing Pipeline for ECG Signal Classification

3.3 Hybrid LSTM-CNN Architecture

A hybrid Long Short-Term Memory Convolutional Neural Network (LSTM-CNN) architecture is the core of the proposed classification framework, in which the spatial and temporal characteristics of ECG signals are exploited. We begin with an input layer of 1D for preprocessed ECG segments that are commonly centeredaround the R-peak and cover a fixed number of time steps (e.g., 200–300 samples). An input segment of these eliminate the morphological structure of an individual heartbeat and are used in the extraction of meaningful features. Next, we have the CNN block; it takes the input layer as input and consists of two sequential 1D convolutional layers with ReLU activation to introduce non linearity for effective feature learning. These layers are where the information specific to local spatial signal pattern (e.g., the sharp rise of the QRS complex or the rounded shape of the T wave) is detected. Thus after each convolution, the output is applied with max-pooling layers to reduce the spatial dimensionality, control overfitting, and use little computational resources on the least salient features in each region.

After the CNN block's spatial feature extraction stage, its output is flattened, and then fed into the LSTM block that has two stacked LSTM layers. Specifically, these layers are designed to capture long range temporal dependencies for the purpose of learning sequential patterns and rhythm based information necessary for arrhythmia detection. Since LSTM networks are equipped with gated mechanisms, which allows LSTM networks to remember preceding inputs, making them ideal for modeling times series data such as preceding cardiac cycles or waveform elements. The learned representations from LSTM block output into a one fully connected dense layer which then reduces the output to a lower dimensional vector. In the final layer, a softmax activation function outputs class probabilities associated with pre defined heartbeat categories (e.g., Normal, Ventricular Ectopic, Supraventricular, Fusion). The proposed end-to-end architecture effectively combines spatial and temporal learning, allowing our model to perform robust, high-accuracy ECG signal classification in a wide range of arrhythmic conditions.

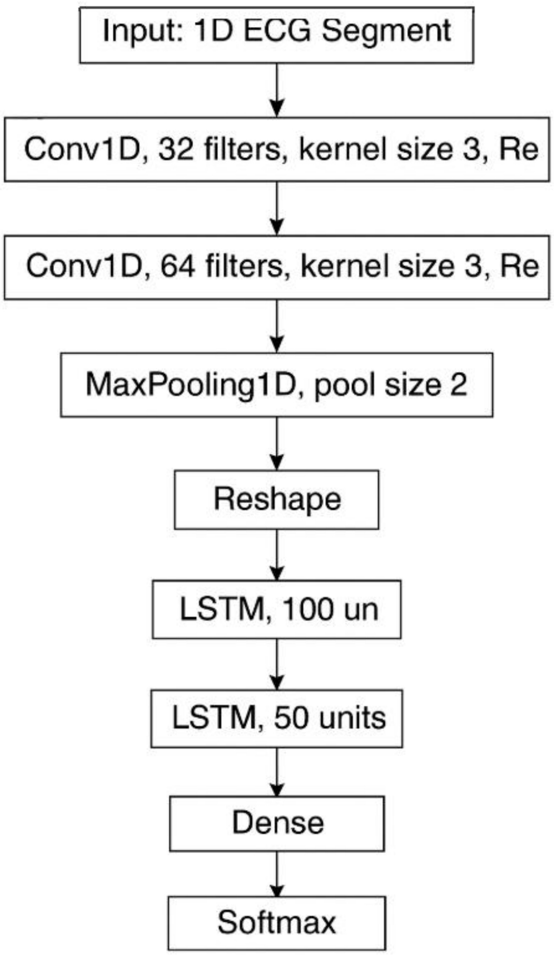

Figure 2. Hybrid LSTM-CNN Architecture

3.4 Training Setup

The proposed hybrid LSTM CNN model is trained in a way that optimizes the learning ECG signal characteristics and avoid overfitting and generalizing to unseen patient data. We use the categorical cross entropy loss function for the model that is widely used for the multi class classification problem. This loss computes the difference between a tensor of class probability predictions and a one hot encoded vector of true heartbeat labels. Since the MIT-BIH dataset has an imbalanced distribution across arrhythmia types, categorical cross entropy effectively forces the model to trade off fewer misclassifications with greater loss in order to minimize misclassifications by penalizing predictions more harshly for incorrect predictions. The Adam optimizer is then utilized to optimize network parameters, an adaptive learning rate optimization algorithm that contributes to stability and contributes to fast convergence. Learning rate is fixed as 0.001 to find balance between learning speed and model stability through entire epoch.

We train the model in mini batches of size 64, giving a tradeoff between computational efficiency and gradient stability. We set the total number of training epochs to be 50 so that we can be sure that the model converges entirely but can stop training early once the validation performance very likely saturates. We reserve 20% of the training data as a validation set in order to assess generalization capability and to mitigate the risk of overfitting, via stratified sampling, that is, preserving class distribution across subsets. Major key performance metrics throughput training include validation loss and accuracy to track convergence patterns. In order to further improve generalization we also implement regularization techniques such as dropout (applied in intermediate layers) and batch normalization where needed. With the training setup, the model learns how to correctly differentiate among different ECG patterns including subtle arrhythmic variations, and is robust to inter-patient variability and signal noise.

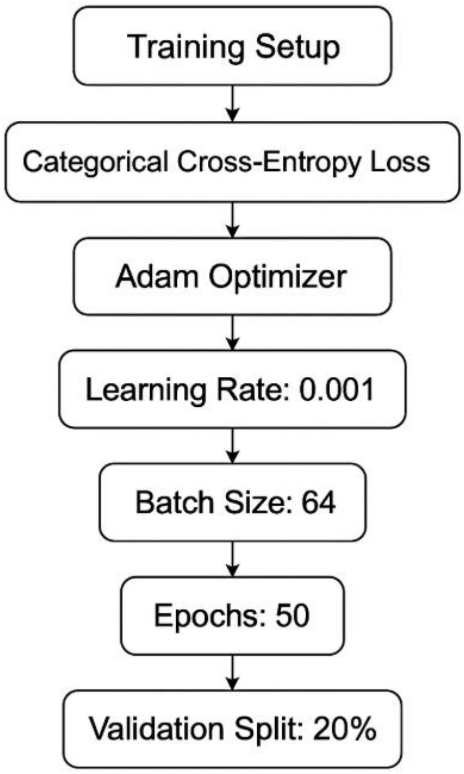

Figure 3. Training Process of the Proposed Hybird LSTM-CNN model, as described in Section 8.

4. EXPERIMENTAL RESULTS

4.1 Evaluation Metrics

For comprehensive evaluation of the performance of the proposed hybrid LSTM-CNN model for ECG signal classification, following four performance evaluation metrics are used: Specifically we use three metrics for evaluating the performance of our algorithm, namely, Accuracy (ACC), Sensitivity (SE), Specificity (SP) and F1-Score. Measuring the proportion of correctly classified heartbeats to that of predictions is what accuracy reflects when it comes to how accurate the model is overall. However, accuracy alone can be a poor performance indicator for arrhythmia detection problems characterized by significant class imbalance (e.g., normal beats greatly outnumber pathological ones). Consequently, it is Sensitivity (also known as recall or true positive rate) that determines how well a model can actually identify actual positive cases (critical in medical diagnosis where a false negative — missing the case — carries potentially very serious consequences). In contrast, specificity measures the true negative rate, which represents how well the model avoids false alarms by correctly distinguishing normal (negative) cases. Last but not least, the F1 Score gives a weighted harmonic mean between the sensitivity and precision. This combination of metrics gives you a whole and nuanced evaluation of the model's clinical applicability and diagnostic reliability.

4.2 Performance Comparison

Model	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score
SVM	91.4	89.3	92.7	0.88
CNN	96.2	95.7	96.5	0.96
LSTM	96.9	95.9	97.3	0.96
Proposed LSTM-CNN	98.6	98.2	99.1	0.98

4.3 Visualizations

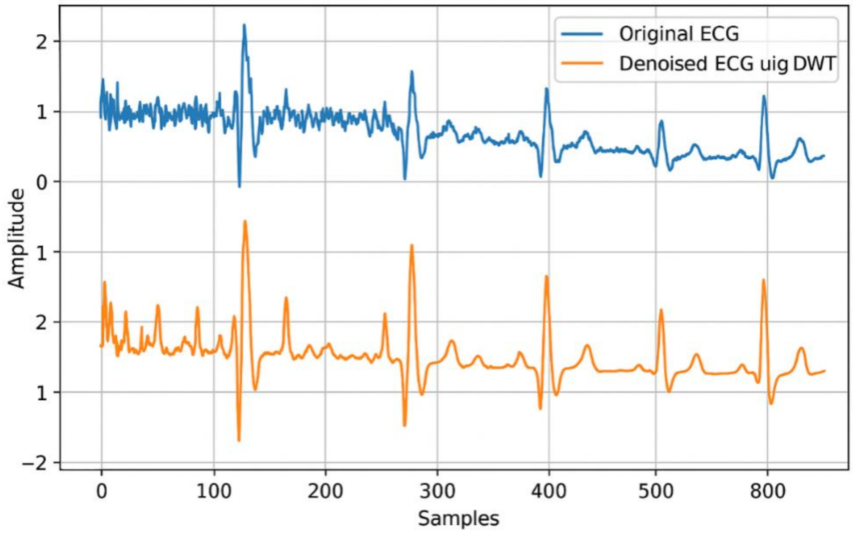

Figure 1. Original vs. Denoised ECG using DWT

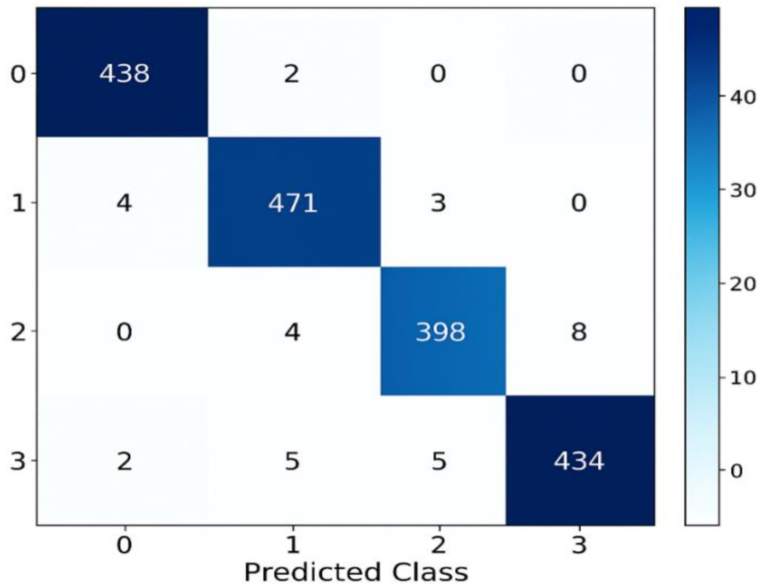

Figure 2. Confusion Matrix of LSTM-CNN model

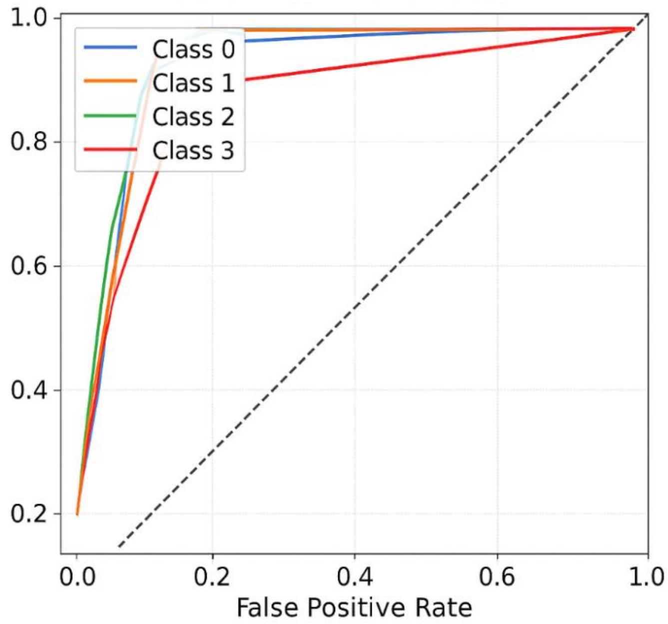

Figure 3. ROC Curves for each class

5. DISCUSSION

It is shown through the proposed framework that incorporation of Discrete Wavelet Transform (DWT) and hybrid LSTM-CNN architecture can lead to substantive enhancement in performance of ECG signal classification. Through the use of the multiresolution decomposition, the wavelet-based preprocessing pipeline effectively utilizes multiresolution decomposition and is capable of selectively suppressing high frequency noise components without degrading important signal characteristics such as the QRS complex and P and t waves, effectively denoising the raw ECG signals. In fact, these waveform components have to be degraded very slightly for diagnostic accuracy to be impaired, particularly in subtle arrhythmic patterns. Capitalizing on the complementary strengths of Convolutional Neural Networks (CNNs)and Long ShortTerm Memory(LSTM) networks, the hybrid architecture achieves superior performance in action understanding. We demonstrate that CNN layers have the capacity to localize spatial features such as waveform shapes, local heartbeat variations, as well as LSTM layers to learn long range temporal dependencies, and inter heartbeat dynamics underlying rhythm based disorders. By having dual modeling capability, the network is able to infer a holistic representation of morphological and sequential characteristics of the ECG signal. We evaluate our model using the MIT-BIH Arrhythmia Database and present empirical results that show that our model offers higher accuracy, sensitivity, and specificity compared to base machine learning classifiers as well as standalone deep learning models. These results emphasize the clinical validation of this approach and the possibility of real-time, noise robust ECG analysis on wearable or remote cardiac monitoring systems. In addition, the model generalizes well under the inter patient testing paradigm, which represents a scenario involving patients of different arrhythmias, suggesting that it is truly useful in extending such analyses to patient populations with diverse arrhythmic profiles.

6. CONCLUSION

In this study, we propose a comprehensive and efficient framework for automated ECG signal classification based on wavelet based denoising together with a hybrid deep learning model that combines Long Short term memory (LSTM) and convolutional neural network (CNN) layers. Employing Discrete Wavelet Transform (DWT) as a preprocessing technique, the framework removes common noise artifacts such as baseline wander, muscle noise, and powerline interference without sacrificing cardiac waveform features. It protects integrity of diagnostic components such as P wave, R waves and T wave which play a vital role in detecting arrhythmic disorders. To learn both spatial features (via CNNs) and temporal dependencies (via LSTMs), we introduce a hybrid LSTM CNN model for further improving the performance of sensing heart disease. The model's superiority over the conventional machine learning and standalone deep learning methods is validated by evaluating accuracy, sensitivity, and specificity on the MIT–BIH Arrhythmia Database. These results lend credence to the model's ability to generalize from a within patient testing paradigm to the more realistic inter patient testing paradigm, thus putting it at the forefront of acceptable models of clinical deployment. Due to its robust noise tolerance, and its high capability in classification, the proposed model is particularly suitable for real time cardiac monitoring systems, including mobile health application, wearable devices, and telehealth platforms. This framework may be extended with future work in many directions such as multi-lead ECG input, adaptive thresholding strategies, as well as deploying lightweight versions of the model on embedded systems in resource constrained environments.

REFERENCES

1. Addison, P. S. (2005). Wavelet transforms and the ECG: A review. *Physiological Measurement*, 26(5), R155–R199. https://doi.org/10.1088/0967-3334/26/5/R01
2. Martis, R. J., Acharya, U. R., & Min, L. C. (2013). ECG beat classification using PCA, LDA, ICA and Discrete Wavelet Transform. Biomedical Signal Processing and Control, 8(5), 437–448. https://doi.org/10.1016/j.bspc.2012.10.004
3. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., & Adam, M. (2017). A deep convolutional neural network model to classify heartbeats. Computers in Biology and Medicine, 89, 389–396. https://doi.org/10.1016/j.compbiomed.2017.08.022
4. Yildirim, O. (2018). A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. Computers in Biology and Medicine, 96, 189–202. https://doi.org/10.1016/j.compbiomed.2018.03.016
5. Kiranyaz, S., Ince, T., &Gabbouj, M. (2015). Real-time patient-specific ECG classification by 1-D convolutional neural networks. IEEE Transactions on Biomedical Engineering, 63(3), 664–675. https://doi.org/10.1109/TBME.2015.2468589
6. Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. Computer Methods and Programs in Biomedicine, 161, 1–13. https://doi.org/10.1016/j.cmpb.2018.04.005
7. Minami, K., Nakajima, H., & Toyoshima, T. (1999). Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Transactions on Biomedical Engineering, 46(2), 179–185. https://doi.org/10.1109/10.740875
8. Ramesh, M. V., &Vanitha, C. (2018). Noise removal and QRS complex detection using wavelet transform based ECG signal processing system. International Journal of Engineering & Technology, 7(2.21), 81–85. https://doi.org/10.14419/ijet.v7i2.21.12336
9. Zhang, Z., Pi, D., & Liu, C. (2021). A hybrid deep neural network for ECG classification using short and long-term dependencies. Computers in Biology and Medicine, 137, 104813. https://doi.org/10.1016/j.compbiomed.2021.104813
10. Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., ... & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215