Journal of Information Systems Engineering and Management
2025, 10(21s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Cardiovascular Disease Detection from Multisource, Multilabel ECG Signalswith Wavelet Scattering Transform, Time-Distributed CNNs, and RNNs

Ankur Rana[1*], Vivek Kumar[2]
[1,2] Department of Computer Science and Engineering,Quantum University, Roorkee-247667, (Uttarakhand), India
*Corresponding author: Ankur Rana, ankur.rana87@gmail.com
profvivekkumar@gmail.com

ARTICLE INFO
Received: 18 Dec 2024
Revised: 02 Feb 2025
Accepted: 14 Feb 2025

ABSTRACT
The classification of cardiac-abnormality patterns plays a crucial role in the diagnosis and treatment of cardiovascular diseases. With the advent of deep learning techniques, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), there has been a significant advancement in the accuracy and efficiency of cardiovascular disease classification with electrocardiogram (ECG) data. However, with the availability of multitudes of freely available multi-source ECG data today, more attempts are required to develop new models that can handle and perform ell on these datasets simultaneously. In this study, an attempt is made to develop a novel deep learningclassification model with multi-source and multi-label ECG dataset for cardiovascular disease classification. Since ECG signals from multiple sources are of different lengths, standardization is done using wavelet scattering transform. Wavelet scattering transform provides time-frequency features that are independent of length of an ECG signal. A deep learning model is then used to perform cardiovascular disease detection and classification.The model is ahybrid synergistic architecture that uses CNNs in time-distributed fashion and RNNs in bidirectional many-to-many fashionThe CNNs exploit inter-lead correlations and provide temporally compressed features.These temporally compressed featuresare exploited via RNNs to provide a well-generalized model. The multi-source ECG dataset used here is composed of 4 popular 12-lead multi-label ECG datasets available publicly for research purposes.The proposed hybrid model performed satisfactorily overall on a 27-class classification scenario.
Keywords: Multi-source multi-label ECG data, cardiovascular disease, wavelet scattering coefficients, deep learning.

INTRODUCTION

Cardiovascular disease classification using ECG data is essential for early detection and diagnosis of various cardiac abnormalities. According to the World Health Organization, the classification of cardiovascular disease plays a crucial role in the diagnosis and treatment of cardiovascular diseases. The examination of variation in ECG waves can be used to detect several cardiovascular abnormalities.An electrocardiogram (ECG) is a visual depiction of the electrical signals produced by the heart[1]. It is employed to detect and diagnose a range of cardiac conditions and irregularities., An ECG signal consists of P, QRS complex, and T waves as shown in Figure 1[1].

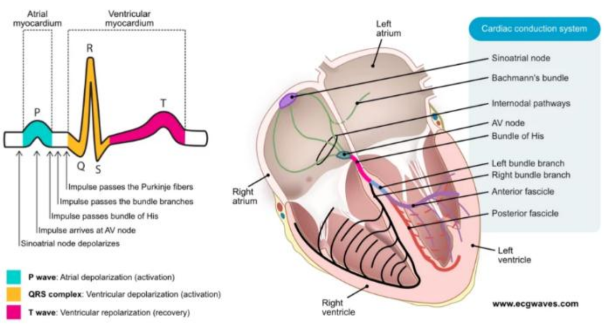


Figure 1 A standard ECG signal decomposition and its correlation to heartbeat progress[2].

A wide variety of heart conditions can be identified by studying the changes in these waves. The extraction of waveforms from the ECG cycle has been the subject of several techniques[1]. Filtering the data, making unique, engaging blocks for each peak, and establishing a constant threshold point are all parts of these techniques. The classification of cardiac abnormality patterns using ECG data is another important goal for determining the heart's health. Authors in [3]notes the need of competency in ECG interpretations and arrythmia management. They realized that hospitals' skilled staff has low competency on ECG interpretation knowledge. Their study recommended use of ECG monitors loaded with automatic detection and diagnosis.

Traditional methods relied on handcrafted features and rule-based algorithms, which were limited in their ability to capture complex patterns present in multi-source ECG signals[4]. Deep learning techniques have emerged as powerful tools for automated feature extraction and classification, enabling more accurate and robust analysis. With the establishment of deep learning techniques, particularly convolutional neural networks (CNNs) and recurrent neural networks (RNNs), there has been a significant advancement in the accuracy and efficiency of cardiovascular disease classification using ECG data[5], [6].

Several reviews and studies have been reported utilizing Recurrent Neural Networks or more commonly termed as RNNs and Convolutional Neural Networks or CNNs in ECG signal analysis. However, in this study, the review is exhaustive and limited to the issue at hand. Thestudy in [5] examines several deep learning methodologies, especially RNNs, for the purpose of detecting arrhythmia in ECG signals. The study explores the utilization of RNN structures, namely long short-term memory (LSTM) and gated recurrent unit (GRU), for the purpose of sequence modelling and classification tasks within the ECG signal analysis domain. The study offers valuable insights into the capabilities and constraints of RNN-based models, and identifies intriguing avenues for future research in this field.The study in [7] introduces an innovative deep learning framework for classifying ECG arrhythmias. The framework combines a 2-D CNNs with an attention mechanism and bidirectional LSTM layers. The attention method enables the model to selectively concentrate on significant portions of the input ECG signals, while the bidirectional LSTM successfully captures temporal relationships or dependencies. Theirproposed model demonstrates a high level of performance in accurately classifying arrhythmias on widely recognized datasets such as the MIT-BIH, and the CPSC.The study in [8] investigates the utilization of transfer learning from pre-trained 2-D CNNs in conjunction with bidirectional LSTM layers for the purpose of classifying ECG arrhythmia patterns. Their work showcases the efficacy of synergistic use of CNNs and RNNs in utilizing information from extensive picture datasets to enhance the accuracy of ECG classification tasks. Their proposed model also demonstrates comparable performance on established ECG classification benchmark datasets while requiring less time and computer resources for training.The research presented in [9]suggested model that combines a CNNs with a transformer to categorize ECG signals. The model utilizes a window function to partition the ECG data into several ECG segments. The CNN's feature information retains its temporal information. Their study claimed that the integration of the CNN and enhanced transformer ultimately attained an F1 score of 78.6%, on the dataset published in [10] in 2018 offering valuable decision support to doctors and cardiologists.The study in [11] introduces a new deep neural network,*ECGDETR*, which is designed to identify arrhythmia in continuous ECG segments from a single lead. Their suggested technique achieves comparable performance in heartbeat placement and classification when compared to prior efforts. They leveraged inter-heartbeat dependence. Furthermore, their model utilizes a more concise inference approach since it does not need explicit segmentation of heartbeats. A robust and effective unsupervised anomaly detection model in time series ECG data is presented in [12]. Their strategy for detecting anomalies in ECG time series is based on a two-stage sequence prediction method. Their model design consists of an encoder network and a fully-connected decoder network. Their proposed model identified abnormalities in human cardiac time series signals, including premature ventricular contractions (PVC), supraventricular premature (SP), and other ECG anomalies. Their findings from analysis of the MIT-BIH and the ECG5000 Arrhythmia datasets indicated towards the effectiveness of encoder-inspired CNNsin replacing the existing method for the purpose of ECG signal analysis and classification. They gained success in establishing that their model surpassed the other models in terms of various performance metrics.

However, recently, multitudes of ECG data are made available for algorithm development and analysis. The datasets are available with varied lead-counts and signal lengths.These variations in ECG data characteristics makes it difficult for deep learning algorithms to provide high performance over each of these datasets. Also, there are presence of multiple heart conditions reflecting on the ECG. In light of these issues, handling multi-source, multi-labelECG data is a challenge and developing more sophisticated deep learning models that can perform well on such datasets is always beneficial.From the afore-mentioned literature, the following observations are useful in development of deep learning models for multi-source and multi-label ECG data.

- CNNs, RNNs, and Transformer models have been successful in arrythmia detection and classification.
- Synergistic use of CNNs and RNNs in ECG signal analysis is significant and beneficial.
- Breaking longer time-length ECG signals into shorter segments helps improve detection and classification tasks.
- More focus is required on the development of generalized models that can handle and perform well on multi-source and multi-label ECG data.
- Features independent of ECG source-type needs to be investigated for development of generalized DNN models so they can be applicable to all ECG sources i.e. single lead or multi- lead, sampling frequency.

Therefore, in this study, hybridDNN modelbased on CNNs and RNNs is proposed, developed and utilized to perform cardiovascular disease classification in multi-source, multi-label ECG data. The reason for a hybrid architecture is because a simple RNN displays a strong inductive bias towards learning temporally compressed representations whereas the upside of a compressed representation of a sequence is that it is beneficial for generalization which is desirable.In contrast, the CNNs are good feature extractors. However its use in high-diemnsional signals increases computational complexity. Also, keeping the temporal informationintcat while extrcating features via CNNs is trick. A time-distributed variant of the CNNs can in fact extrcat features from data without hampering the timporal information. The temporally compressed feaures can then be exploited by the RNN. The RNN is employed in bidirectional, many-to-many fashion to maximize the temporal information exploitation.This strategy will help develop efficient approach/model in multi-source ECG data based cardiovascular classification. The paper is divided into sections. Section 1 provides premise to the study and relevant literature background. Section 2 and its sub-sections incorporates the materials and methods utilized in the study. Section 3 provides results obtained and its discussion. Section 4 concludes the study.

MATERIALS AND METHODS

2.1 Dataset: Description and Preparation: The 12-Lead multi-source multi-label ECG Database, Statistics, and Preprocessing

Overall, the 12-Lead multi-source multi-label ECG database considered here for study is composed of 4 different databases. These are

1. The PTB and PTB-XL database
2. Georgia 12-Lead ECG Challenge Database
3. The St. Petersburg database i.e. INCART database
4. The China Physiological Signal (CPS) database

Table 1 provides a summary of the individual databases. The multi-source ECG data includes a total of 27 cardiac abnormalities, which are detailed with their abbreviations in Table 2. The sample proportions for these abnormalities in the combined dataset are illustrated in Figure 2. It is clear from Figure2, the 'Sinus Rhythm' (SNR) heart condition accounts for the highest number of samples, while the 'Premature Ventricular Contractions' (PVC) heart condition has the lowest count. Table 2 fatherprovides a more detailed breakdown of sample proportions for each individual database.

Table 1 Properties of the different sources of 12-lead ECG database to be used in the study.

Database Attribute	PTB and PTB-XL	Georgia	CPS	INCART
Subjects	290, 18885	15742	9458	32
Records (length in seconds)	516, 21,837 (NA)	20,678 (10)	13,256 (6-60)	74 (NA)
Sampling rate (Hz)	1000 (PTB), 500 (PTB-XL)	500	500	256
Mean Age	56.3, 59.8	60.5	61.1	56
Male, n (%)	211 (73), 10197 (54)	8500(54)	5013 (53.1)	17(54)
Leads, n	12	12	12	12
Classes	9, 71	24	23	10

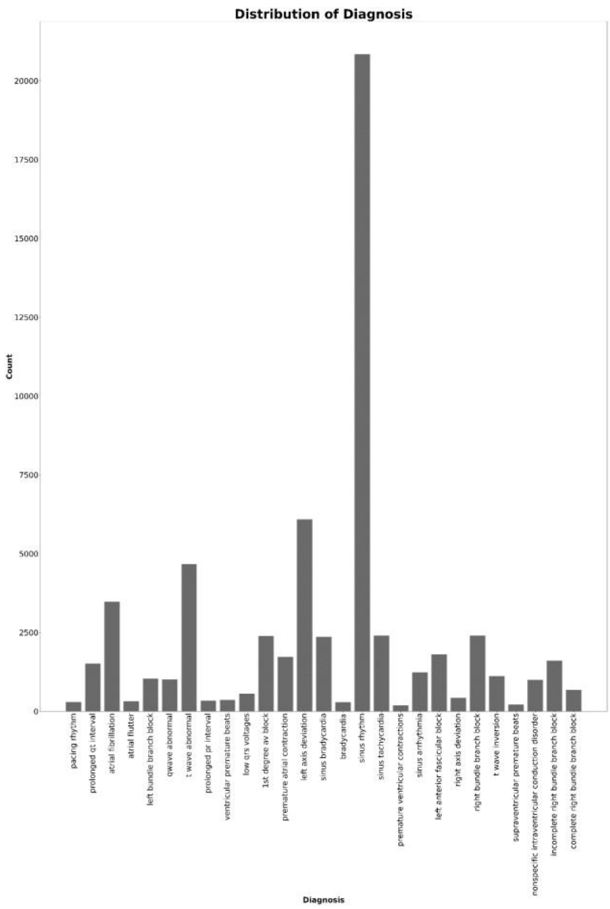


Figure 2 Bar chart showing the class-wise distribution of data samples.

Table 2 Summary of ECG data diagnosis sample proportions in individual database.

Diagnosis	Abbreviation (Total samples)	Sample proportions					
Complete right bundle	CRBBB (701)	28	0	542	113	0	18

branch block							
Atrial utter	AFL (423)	186	1	73	54	0	109
Left anterior fascicular block	LAnFB (1916)	180	0	1626	0	0	110
Prolonged PR interval	LPR (340)	0	0	340	0	0	0
Premature ventricular contractions	PVC (188)	0	0	0	188	0	0
Supraventricular premature beats	SVPB (215)	1	0	157	53	4	0
Right axis deviation	RAD (465)	83	0	343	10	0	38
Nonspecific intraventricular conduction disorder	NSIVCB (1093)	203	0	789	4	1	96
Ventricular premature beats	VPB (543)	357	0	0	8	0	178
Left bundle branch block	LBBB (1197)	231	0	536	274	0	156
Low QRS voltages	LQRSV (748)	374	0	182	0	0	192
Bradycardia	Brady (289)	6	0	0	271	11	1
Pacing rhythm	PR (301)	0	0	296	30	0	2
Q wave abnormal	QAb (1252)	464	0	548	1	0	239
Sinus arrhythmia	SA (1476)	455	0	772	11	2	236
Left axis deviation	LAD (6564)	940	0	5146	0	0	478
Premature atrial contraction	PAC (2188)	639	0	398	689	3	459
T wave inversion	TInv (1550)	812	0	294	5	1	438
Incomplete right bundle branch block	IRBBB (1871)	407	0	1118	86	0	206
Sinus tachycardia	STach (3050)	1261	1	826	303	11	648
Prolonged QT interval	LQT (2253)	1391	0	118	4	0	740
Right bundle branch block	RBBB (3051)	542	0	0	1858	2	649
Atrial defibrillation	AF (4026)	571	14	1514	1274	2	552
1st degree AV block	IAVB (2946)	769	0	797	828	0	552
Sinus bradycardia	SB (3219)	1677	0	637	45	0	860
T wave abnormal	TAb (5792)	2306	0	2345	22	0	1119
Sinus rhythm	SNR (21944)	1752	78	18092	922	0	1100
Dataset		Georgia (10344)	PTB (490)	PTB-XL (21837)	CPSC (10330)	INCART (74)	Validation (6630)

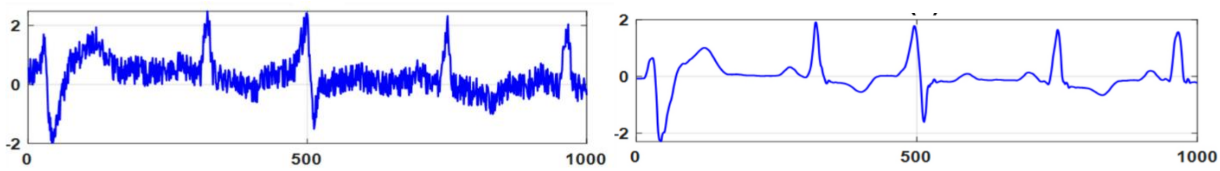
2.2 Data Preparation

The inherent noise in the recordings and the imbalance in cardiac-abnormality sample proportion are the issues that need attention and are addressed in this section.

Signal Denoising and Filtering

ECG signals often suffer from various types of noise, including baseline wander (BW), power-line interference (PLI), motion artifacts, and physiological artifacts. Among these, BW and PLI are the most significant contributors to signal degradation, affecting both visual interpretation and automated diagnostics. Additionally, the frequency characteristics of ECG signals change over time, making noise removal essential for accurate analysis. To address this, preprocessing techniques are applied to eliminate artifacts and noise, ensuring the ECG data is suitable for reliable interpretation.[13], [14], [15], [16], [17].

In this study, a denoised ECG signal is produced by addressing various types of noise, including electrode contact noise, power-line interference, muscle contractions, motion artefacts, baseline wander, and random noise[16].A step-by-step noise reduction process is applied to clean the raw ECG signals.Figure 3(a) illustrates a sample ECG signal affected by these noises. The noise reduction process begins with a Butterworth low-pass filter to eliminate frequency components above 50 Hz. Next, a LOESS smoother is used to minimize the effects of baseline wander. Finally, a non-local mean algorithm is applied to handle remaining noise. The resulting denoised ECG signal after preprocessing is shown in Figure 3(b).



(a) (b)

Figure 3 (a) Raw observed ECG signal, and (b) Filtered and denoised signal.

Imbalanced data

A significant challenge with multi-source databases is the imbalance in sample proportions, which can negatively impact the performance of machine learning algorithms [18]. The imbalance in the multi-source ECG dataset considered here cab be observed from Figure2. To address this, an empirical sample-weight allocation strategy is implemented. We calculate a weight for each of the 27 diagnoses to balance the data, ensuring that the algorithm learns effectively from all labels. The weight values for each diagnosis are provided in Table 3. This balanced and filtered dataset is then used for algorithm development in this study.

Table 3 ECG diagnosis label weight values.

Class	Weight value	Class	Weight value
5	21.27	19	11.93
8	59.04	22	19.38
9	38.76	23	100.23
10	9.00	24	21.62
11	12.46	25	13.38
6	4.61	20	50.47
7	63.38	21	8.97
0	72.08	14	74.83
1	14.24	15	1.03
2	6.20	16	8.97

3	68.63	17	114.63
4	20.70	18	17.38
12	3.54	26	31.55
13	9.14		

2.3 Wavelet Scattering Transforms and Wavelet Coefficients

The Wavelet Scattering Transform is a versatile method used in signal processing and machine learning to extract meaningful features from signals and images. It is designed to handle transformations like shifts (translations) and remain stable under small changes (deformations). Authors in [19] reported the use of wavelet scattering features in emotion recognition with ECG signals. They established the importance of wavelet scattering features as time-frequency features for efficient classification of emotions in humans. By blending the strengths of wavelet transforms and principles inspired by deep learning, it generates reliable representations that are especially effective for tasks such as classification, pattern recognition, and feature extraction.

Wavelet Transform

- Wavelet transforms are a mathematical technique used to break down signals into various frequency components, offering insights into both time and frequency simultaneously.
- Unlike the Fourier Transform, which focuses solely on frequency, wavelets can track how the frequency content of a signal changes over time.
- This approach relies on wavelet functions, which are designed to be localized in both time and frequency domains.

Scattering Transform

- The scattering transform, developed by Stéphane Mallat and his team[20], provides a hierarchical and multi-scale way to represent signals.
- It works by applying a sequence of wavelet transforms, followed by non-linear modulus operations and smoothing through averaging.
- This method resembles the functionality of convolutional neural networks (CNNs) but uses predefined wavelet filters instead of filters learned from data.

Wavelet Scattering Transform Process

Wavelet Decomposition:

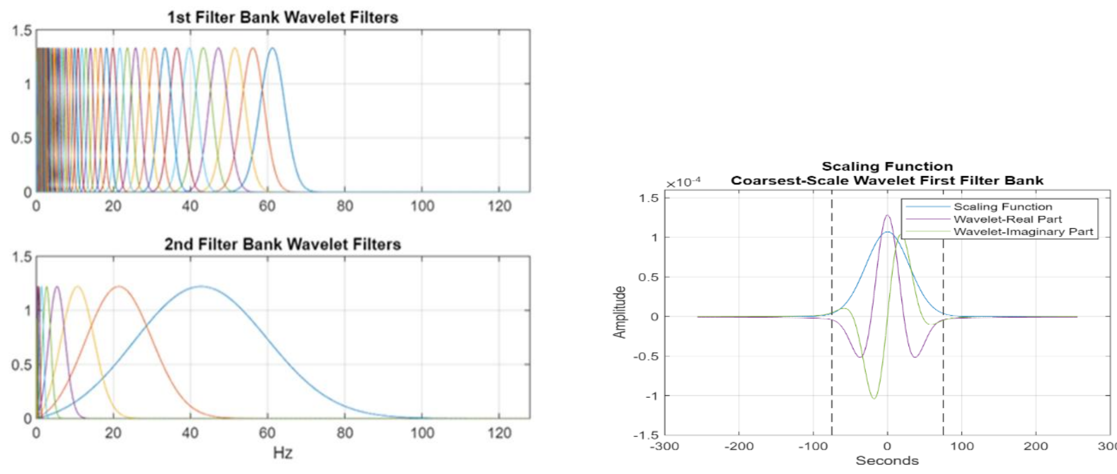
- The signal or image is broken down using multiple wavelet transforms across various scales and orientations.
- This step generates coefficients that capture local changes at different scales and directions.

Non-linear Modulus Operation

- A non-linear modulus operation is applied to these coefficients after each wavelet transform.
- This operation keeps the magnitude of variations while discarding phase information, making the representation more resilient to small deformations.

Averaging

- The modulus coefficients are averaged over local regions to enhance stability and reduce sensitivity to noise and minor deformations.
- This step is key to creating representations that are both translation-invariant and deformation-stable.



(a) (b)

Figure 4 (a) First (top) and second (bottom) filter banks for proposed scattering transform and, (b)Scaling function.

Wavelet Scattering feature computation

The main parameters to configure in a wavelet time scattering network include the time-invariant scale, the number of wavelet transforms, and the number of wavelets per octave within each wavelet filter bank as listed in Table 4. The specific settings for these parameters are also detailed in the Table 4. Figure 4(a) shows the filter banks from the transform used here whereas the scaling function is shown in Figure 4(b).

Table 4 Summary of parameter settings for the wavelet scattering transform used in the study

Parameter	**Setting**
Total number of samples, T	5000
Invariance scale in samples, 2^j	256
Number of wavelets per octave	8

- The input signal is processed using a scattering transform to break it into various frequency components.
- This transform operates across multiple scales, allowing it to capture frequency details at different resolutions.
- The modulus outputs from the first-order scattering are further processed with a second wavelet transform, breaking down the magnitude coefficients to capture higher-order statistical information.
- The resulting scattering coefficients serve as feature representations of the signal, which can then be used in standard machine learning models for classification tasks.

The filters are defined for a certain support size T. In turn, T or the support size corresponds to the size of the input signal. It is important to note here that T is employed under a constraint that itapproximates toa power of 2. In this study, T is set to 2^{19} (4864).Two layers of filter banks are considered here. The number of wavelets per octave in these first-and second-order filter banks are controlled by a parameter Q. In general, the number of non-negligible oscillations in time is proportional to Q.The smaller the value of Q, the broader these filters in the frequency domain are and the narrower they are in the time domain. For ECG signals, a large value for Q_1(between 4 and 16) is often beneficial.Also, the ECG signals are often periodic andare better localized in the frequency domain than they are in time domain. The parameter J specifies the maximum scale of the low-pass filters as a power of 2. In other words, the largest filter will be concentrated in a time interval of size 2^j.The time-length selected for employing the scattering transform is 1-second. A total of 19 such segments are made. A total of 234 scattering coefficients (truncated) to represent each segment. Figure 5 show scattering coefficients obtained post transform. Figure 5(a) show sample signals for 'AF', 'CRBBB', and 'SNR' classes whereas Figure 5(b) shows corresponding zeroth-order, first-order, and second-order coefficients respectively from top to bottom.

Table 5Summary of scattering coefficients for sample records.

Parameter	**Value**		
Sample id	10000	10500	20000
Channel id	10	5	1
label	'AF'	'CRBBB'	'SNR'
T	4864 (~5000)	4864	4864
J	8	8	8
Q	8	8	8
Resulting time steps	4864/256 = 19	4864/256 = 19	4864/256 = 19
Scattering coefficients	234	234	234

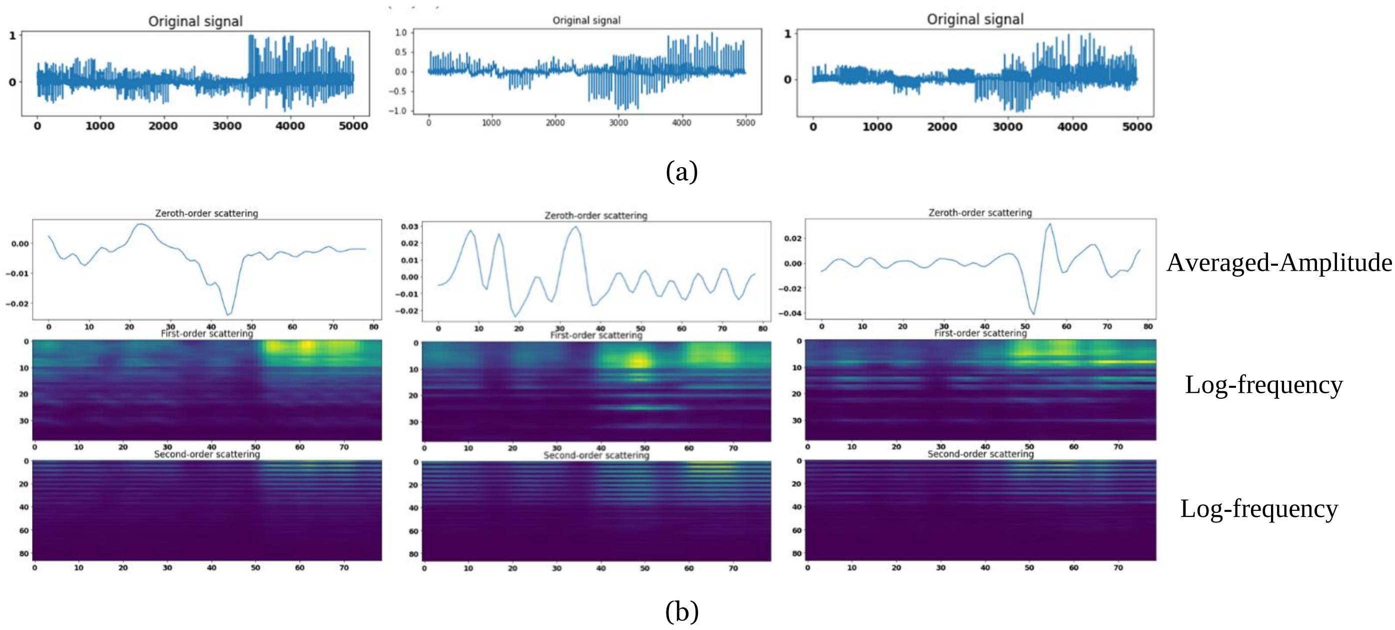


Figure 5 (a) Original signal; left: 'AF', middle: 'CRBBB', right: 'SNR', (b) Average amplitude, first-order scattering coefficients, and second-order coefficients (top to bottom) for left: 'AF', middle: 'CRBBB', and right: 'SNR'.

2.4 Hybrid CNN-RNN model for ECG Signal classification

The study presents a hybrid convolution-recurrent network for cardiovascular disease detection. The convolutional layers exploiting the scattering coefficients is employed in a time-distributed fashion. The recurrent layer in cascade with the convolutional layer extracts the temporal information form the data.CNNs are known for their ability to capture contextual dependencies in data, making them suitable for tasks where extracting patterns are important such as the case in hand in this study. CNNs excel in tasks such as ECG signal analysis, where interpreting context is critical for performance. However, since CNNs limit to extract temporal dependencies due to no-memory architecture.Meanwhile, RNNs, particularly LSTMs and GRUs, are suitable for capturing temporal dependencies in ECG data.

To understand how the hybrid CNN-RNN model acts on the input data, it is important to interpret the input data. Figure 6 sow the architecture of the hybrid model along with transformation of the input data structure. According to Figure 6, the input is a three-dimensional tensor. The first dimension is the time dimension. The entire signal is broken into 19 segments of equal time-length. The second dimension represents number of channels. Since the

ECG signals are recorded via 12 leads across the human body. Therefore, this dimension encapsulates a 12-dimensional vector. Finally, the last dimension of the 3D tensor represents the number of scattering coefficients. Here, 234 coefficients are used to represent one ECG-lead signal.Therefore, thelast dimension is a 234-dimensional vector.A one-dimensional convolution is firstly applied to extract the high-level features from each segmented scattering coefficients-laced signal. This means that one-dimensional convolutions are applied to leaddimension. Note here that each lead signal is already representedvia234 scattering coefficients.The convolution window ofthe first 1D convolutional layer is 3. The structure of the output from this layer is governed by equation 1.

$$n_{out} = \left\lceil \frac{n_{in} + 2p - k}{s} \right\rceil + 1 \tag{1}$$

Where, n_{in} is number of input features, n_{out} is number of output features, k is convolution kernel window size, p is convolution padding size and, s is convolution stride size.The 1D convolution is applied to the lead dimension which is 12D vector. Both padding and the stride size is set to 0. Therefore, accordingto equation 1, the outputs after the first and the second 1D convolutional layer is 9 and 6 respectively. Since all convolutional layers are in time-distributed fashion, the time dimension remains intact. To extract the time information, first, the last tow dimensions are merged to form a feature dimension. An LSTM-based recurrent layer is employed to extract the temporal information. This recurrent layer has two characteristics. One, use of bidirectional flow of activation and two, many-to-many strategy.The bidirectional flow of activation allows to compute the features in both, past-to-future and future-to-past dependence scenarios. Since, this study addresses a detection task in anon real-timescenario, features are allowed to be computed in both directions. The many-input to many-output strategy allows to take output from any time-stamp of the recurrent layer. This also helps find out which time-stamp is most relent for efficient detection. Features extracted post the recurrent layer are fed into a fully-connected layer for scaling and dimensionality reduction. Finally, a classification layer with Sigmoid activation is employed. This layer provides likelihood of 27 classes.

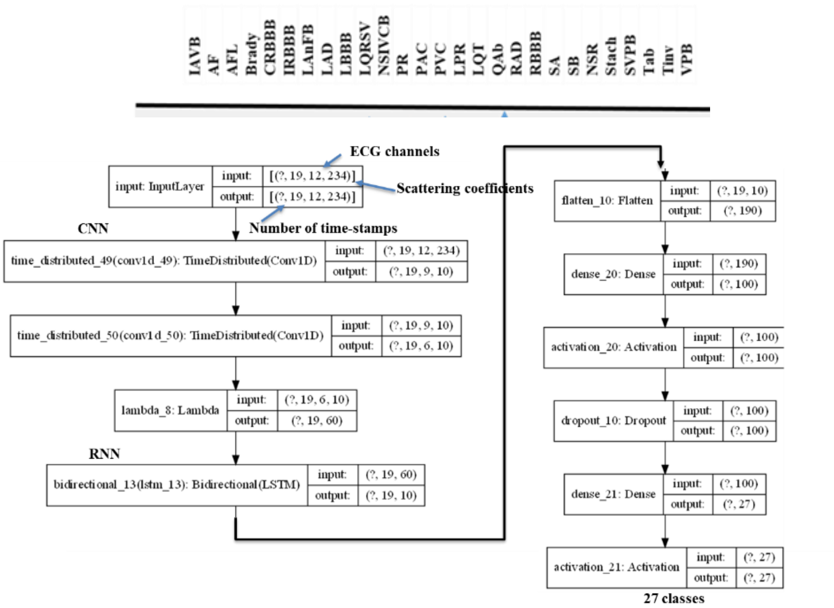


Figure 6 Figure showing the hybrid architecture of proposed hybridCNN-RNN model.

EXPERIMENT SETUP, RESULTS, AND DISCUSSION

2.5 Setup and Model Training

Acardiovascular disease classification model is built. The model architecture summary and hyperparameter settings are listed in Table 6. A total of 7 layers are used. Since the dataset has multi-label ECG records present, a binary cross-entropy is used as loss estimator and Adam as optimizer. A 3-fold cross-validation strategy is also employed to counter bias during training. The sample distribution between model training and model testing is provided in Table 7. With these model training settings, the model is trained for 50 epochs and corresponding training progress is shown in Figure 7.

2.6 Model Performance: results, and Discussion

It is evident from Figure 7 that both training and validation accuracies are around 96% and converging. However, there is a constant bias between the training and the validation accuracy. This could indicate the saturated performance of the proposed model. A quantitative assessment of the model performance on validation and testing samples is listed in Table 8. There is a 3% decline in classification accuracy from validation set to testing set. This could be due to the diverse class sample distribution in the multi-source dataset with 27 classes (refer Figure 1) that are overlapping. A different scoring mechanism would be more appropriate to reflect on the performance of the proposed model. The next section covers such an assessment scenario.

Table 6 The proposed model architecture and hyperparameter settings.

ECG signal classification model with DNN	
Number of layers	7
Feedforward network dropout size	0.2
Number of classes	27
Epochs	50
Learning rate	1e-4
Weight decay rate	1e-4
Cross-validation folds	3
loss	Binary Cross-Entropy
Optimizer	Adam
Performance metric	Accuracy

Table 7 Summary of the multi-source multi-label dataset sample distribution.

Set	Number of Samples
Training and validation	30000
Testing	13000

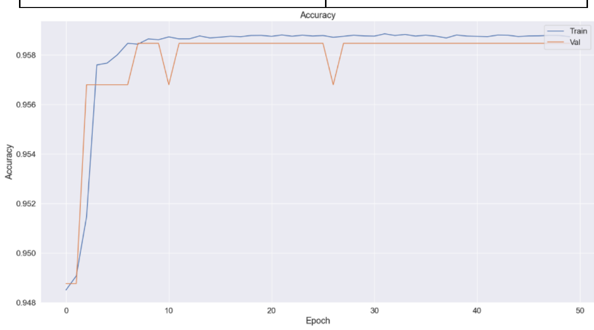


Figure 7Training progress curve for the proposed model.

Table 8 Model performance on validation set of the multi-source multi-label dataset.

Attribute	Overall loss		Overall accuracy	
	Validation	Testing	Validation	Testing
Value	0.14	0.18	95.8%	93.8%

A normalized confusion matrix is reported here. This matrix elements are normalized over the number of total recordings.This is because each recording in the multi-label ECG data could have multiple labels and consequently the proposed model can produce multiple outputs for a single recording. The contribution of each recording is normalized in the scoring metric. This is achieved by dividing by the number of output classes with positive value. For each recording k = {1, 2, ...*n}* in the dataset, let x_k be the set of positive labels i.e. ground truth or actual label and y_k be the set of positive classifier outputs. Therefore, thenormalized confusion matrix is calculated as per equation 2.

$$a_{ij} = \sum_{k=1}^{n} a_{ijk}, \tag{2}$$

Where,

$$a_{ijk} = \begin{cases} \frac{1}{|x_k \cup y_k|}, & \text{if } c_i \in x_k \text{ and } c_j \in y_k, \\ 0, & \text{otherwise.} \end{cases}$$

The quantity $|x_k \cup y_k|$ is the number of distinct classes with a positive label and/or classier output for recording k.

The normalized confusion matrix obtained for the proposed model is shown in Figure 8.The proposed model is able to classify the 'PR' with highest score of 0.80. It is also evident from the matrix is that the normal sinus rhythm or 'SNR', t-wave abnormal or 'TAb', and left axis deviation or 'LAD' classesare mixing with other classes. This is may be due to the large proportion of 'SNR' (21944), 'LAD' (6564) and 'Tab' (5792) samplesin the whole dataset (42,000 approx. (post filtering)) and even after weight allocation, the number of samples are still able to bias the model's performance.This issue could be a scope for further investigation where more optimized nature of weight allocation strategy could be explored or sample imputation can be executed. It seems as if the model is giving a lot of false negatives as well i.e. a cardiovascular disease is present in the signal but the model is identifying the signal as normal. For example, the 'SA' cardiovascular disease is identified as 'SNR' with a significantly high score of 0.42 whereas the score for it being correctly identified is only 0.01. In summary, the model is able to perform satisfactory classification over 27 classes however there is scope for improvements. To compare the performance of the proposed model, the one-dimensional FCN (fully connected network) and one-dimensional Resnet-1D (Residual network) that has been used for ECG signal classification [5], [21] are considered here. Figure 9 and Figure 10 show scoring matrix for FCN and Resnet-1D respectively. Moreover, for Resnet-1D and the proposed model, thereare 5 classes with scores less than or equal to 0.05namely 'QAb', 'SA', 'RAD', 'TInv', and 'NSIVCB. For FCN, there are 7 classes with scores less than 0.05 namely namely 'QAb', 'PAC', 'SA', 'RAD', 'TInv', 'SVPB', and 'NSIVCB' [22], [23].

Although the Resnet-1D and the proposed model are showing equivalent performance, however, in context to false negatives, the proposed model is performing better than Resnet-1D. The average false negative (others to SNR) for Resnet-1D is 0.18 whereas it is 0.14 for the proposed model.These inferences indicate to the bias nature of FCN model towards classes with high sample proportions. The proposed model is better in terms of generalizationhowever Resnet-1D performs equivalently. Therefore, the superiority of the proposed model for multi-source multi-label ECG data of is inclusive [24].

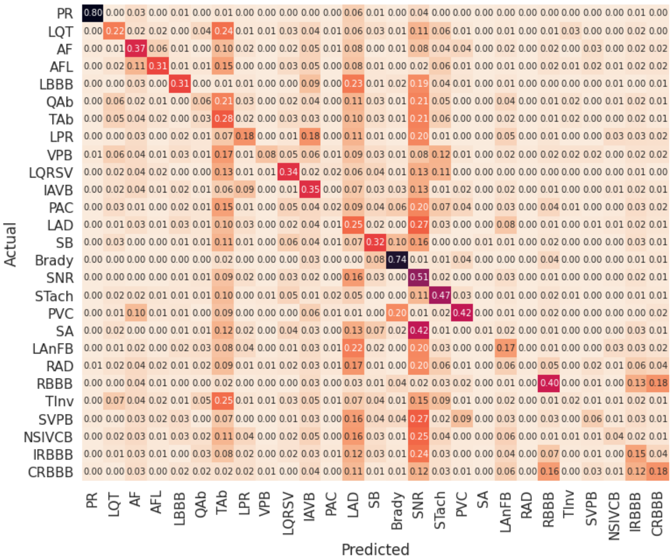


Figure 8 Proposed model performance score matrix.

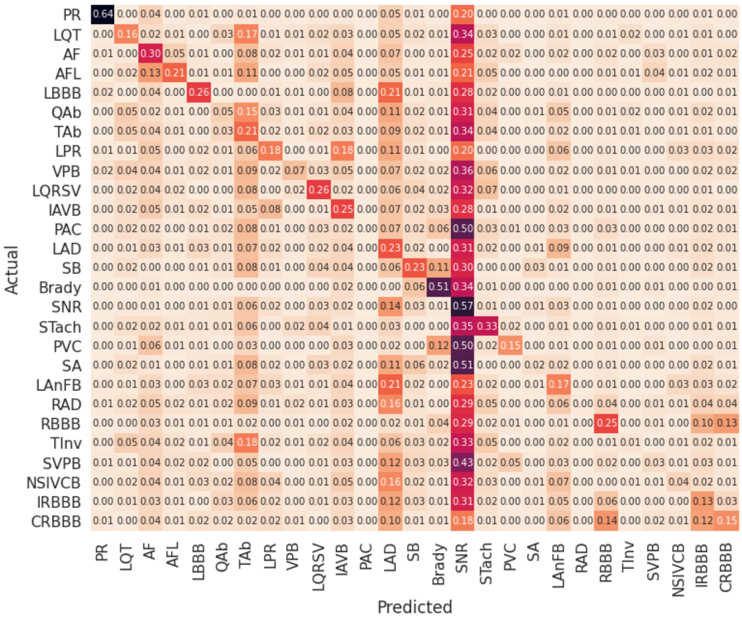


Figure 9Performance score matrix for FCN model.

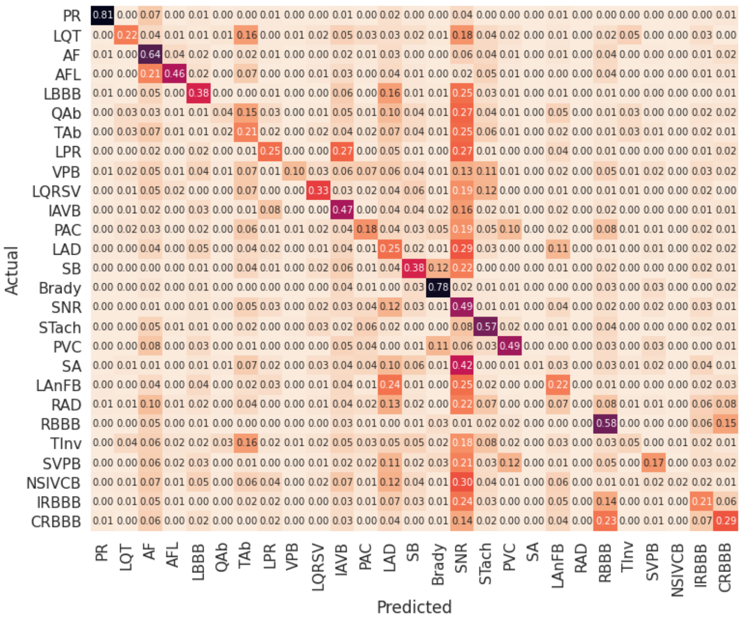


Figure 10 Performance score matrix for Resnet-1D model.

CONCLUSION

In this study, a scattering transform feature extraction approach is utilized which enabled extraction of features based on time-frequency characteristics of ECG observation signals.Scattering coefficients and scalograms features have been identified as potential ECG signal representations. Further, the extracted features are exploited using a hybrid CNN-RNNdeep learning model. The model is composed of one–dimensional time-distributed convolutions and bidirectional recurrent layers for multi-label ECG classification. The time-distributed one-dimensional convolutions along the lead dimension captured the inter-lead correlations whereas the bidirectional many-to-many recurrent layer extracts temporal information. The model is trained multi-source ECG data with 27 classes present in it. The model achieved 96% overall classification accuracy on the test set. Since each ECG record may have more thanone class present.Therefore, a scoring metric is used to evaluate the model performance that reflect a satisfactory performance of the proposed model.However popular models such as Resnet-1D showed comparable performance on the same dataset. A key takeaway from the study is the need for improved strategy for training sample proportion distribution. Especially this study reveals that in a larger class-set (27-classes) classification scenario compounded by overlapping characteristics of classes (cardiovascular diseases), the class-sample distribution plays crucial role.

STATEMENTS AND DECLARATIONS

Competing Interests: Authors are required to disclose financial or non-financial interests that are directly or indirectly related to the work submitted for publication. Please refer to “Competing Interests and Funding” below for more information on how to complete this section.

DATA AVAILABILITY

Data sharing does not apply to this article.

REFERENCES

[1] S. Kaplan Berkaya, A. K. Uysal, E. Sora Gunal, S. Ergin, S. Gunal, and M. B. Gulmezoglu, “A survey on ECG analysis,” May 01, 2018, Elsevier Ltd. doi: 10.1016/j.bspc.2018.03.003.

[2] ECG and ECHO Learning, “Cardiac electrophysiology and ECG interpretation–Cardiovascular Education.” Accessed: Nov. 23, 2024. [Online]. Available: https://ecgwaves.com/topic/introduction-electrocardiography-ecg-book/

[3] M. S. Aljohani, “Competency in ECG interpretation and arrhythmias management among critical care nurses in Saudi arabia: A cross sectional study,” in Healthcare, 2022, p. 2576.

[4] A. Gacek and W. Pedrycz, ECG signal processing, classification and interpretation: a comprehensive framework of computational intelligence. Springer Science & Business Media, 2011.

[5] X. Liu, H. Wang, Z. Li, and L. Qin, "Deep learning in ECG diagnosis: A review," Knowl Based Syst, vol. 227, p. 107187, 2021.

[6] O. Faust, Y. Hagiwara, T. J. Hong, O. S. Lih, and U. R. Acharya, "Deep learning for healthcare applications based on physiological signals: A review," Comput Methods Programs Biomed, vol. 161, pp. 1–13, 2018.

[7] T. Zhang, "Arrhythmias Classification Based on CNN and LSTM_ATTENTION Hybrid Model," in 2021 3rd World Symposium on Artificial Intelligence (WSAI), 2021, pp. 58–63.

[8] M. Salem, S. Taheri, and J. Yuan, "ECG arrhythmia classification using transfer learning from 2-dimensional deep CNN features," in 2018 IEEE biomedical circuits and systems conference (BioCAS), 2018, pp. 1–4.

[9] C. Che, P. Zhang, M. Zhu, Y. Qu, and B. Jin, "Constrained transformer network for ECG signal processing and arrhythmia classification," BMC Med Inform Decis Mak, vol. 21, no. 1, p. 184, 2021.

[10] F. Liu et al., "An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection," J Med Imaging Health Inform, vol. 8, no. 7, pp. 1368–1373, 2018.

[11] R. Hu, J. Chen, and L. Zhou, "A transformer-based deep neural network for arrhythmia detection using continuous ECG signals," Comput Biol Med, vol. 144, p. 105325, 2022.

[12] A. Alamr and A. Artoli, "Unsupervised transformer-based anomaly detection in ECG signals," Algorithms, vol. 16, no. 3, p. 152, 2023.

[13] S. Agrawal and A. Gupta, "Fractal and EMD based removal of baseline wander and powerline interference from ECG signals," Comput Biol Med, vol. 43, no. 11, pp. 1889–1899, 2013.

[14] J. S. Sahambi, S. N. Tandon, and R. K. P. Bhatt, "Quantitative analysis of errors due to power-line interference and base-line drift in detection of onsets and offsets in ECG using wavelets," Med Biol Eng Comput, vol. 35, pp. 747–751, 1997.

[15] W. C. Members et al., "ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction—executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 1999 Guidelines for the Management of Patients With Acute Myocardial Infarction)," Circulation, vol. 110, no. 5, pp. 588–636, 2004.

[16] R. R. Sharma and R. B. Pachori, "Baseline wander and power line interference removal from ECG signals using eigenvalue decomposition," Biomed Signal Process Control, vol. 45, pp. 33–49, 2018.

[17] S. Chatterjee, R. S. Thakur, R. N. Yadav, L. Gupta, and D. K. Raghuvanshi, "Review of noise removal techniques in ECG signals," IET Signal Processing, vol. 14, no. 9, pp. 569–590, 2020.

[18] F. Deng, S. Tu, and L. Xu, "Multi-source unsupervised domain adaptation for ECG classification," in 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2021, pp. 854–859.

[19] A. Sepúlveda, F. Castillo, C. Palma, and M. Rodriguez-Fernandez, "Emotion recognition from ecg signals using wavelet scattering and machine learning," Applied Sciences (Switzerland), vol. 11, no. 11, Jun. 2021, doi: 10.3390/app11114945.

[20] S. Mallat, A wavelet tour of signal processing. Academic Press, 1999.

[21] E. Jing et al., "ECG heartbeat classification based on an improved ResNet-18 model," Comput Math Methods Med, vol. 2021, 2021.

[22] Wasim Khan, Mohammad Ishrat, Mohammad Nadeem Ahmad, Shafiqul Abidin et al, " Enhancing Anomaly Detection in Attributed Networks Using Proximity Preservation and Advanced Embedding Techniques", IEEE Access, ISSN:2169-3536, February 2025.

[23] P. S. Ramesh , P. Srivani , Miroslav Mahdal ,, Lingala Sivaranjani , Shafiqul Abidin et al", Contextual Cluster-Based Glow-Worm Swarm Optimization (GSO) Coupled Wireless Sensor Networks for Smart Cities", Sensors, MDPI, July 2023 , pp 1-26

[24] Shafiqul Abidin, Amit Swami, Edwin Ramirez-Asis, Joseph Alvarado-Tolentino, Rajesh Kumar Maurya, Naziya Hussain "Quantum Cryptography Technique: a way to Improve Security Challenges in Mobile Cloud Computing (MCC)", Materials Today: Proceedings, ISSN: 2214-7853, July 2021