RESEARCH PAPER
Homepage: https://frai.journals.umz.ac.ir/
Future Research in AI & IoT, 2025, 1(2)
DOI:

Detection of Cardiac Arrhythmias Using Transfer Learning and Deep CNN-LSTM Features in the Time–Frequency Domain

Samira Mavaddati[1]*, Hamidreza koohi[2], Ziad Kobti[3]

[1*] *Associate Professor, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran*

[2,3] *School of Computer Science, University of Windsor, Ontario, Canada*

Article Info
Received -----
Accepted -----
Available online -----

Keywords:
Cardiac Arrhythmia Detection, Deep Learning, Transfer Learning, Time–Frequency Analysis, CNN-LSTM, Continuous Wavelet Transform.

Abstract:
Cardiac arrhythmias are a major concern in clinical cardiology, often leading to critical outcomes if not diagnosed promptly. This research introduces a novel deep learning-based architecture that merges time–frequency signal transformation, transfer learning, and sequential modeling to detect arrhythmias from ECG signals. The raw signals are first converted into scalogram representations using continuous wavelet transform (CWT), emphasizing both time and frequency patterns. These images are then passed through a pre-trained convolutional neural network (CNN) to derive abstract, high-level features. To model temporal dependencies in heart rhythms, an LSTM network follows, capturing dynamic signal behavior. Finally, a dense layer classifies the signals into normal and abnormal classes. Evaluations conducted on publicly available ECG datasets show the proposed method achieves 98.76% accuracy, surpassing several current models. These results affirm the effectiveness of integrating transfer learning with temporal modeling in automated arrhythmia detection, paving the way for improved clinical workflows.

***Corresponding Author:** s.mavaddati@umz.ac.ir

Supplementary information: Supplementary information for this article is available at https://frai.journals.umz.ac.ir/

Please cite this paper as:

1. Introduction

Cardiovascular diseases (CVDs) are among the leading causes of death worldwide, with over 17.9 million deaths annually [1]. Early diagnosis of heart disease is crucial for effective treatment and prevention, making it an important problem in the field of healthcare that can lead to more effective treatment and prevention, resulting in a reduction in the burden of heart disease on individuals and society. It also helps decrease the healthcare costs associated with late detection and intervention. Signal processing-based and machine learning-based algorithms are capable of significantly improving the ability to recognize and diagnose heart diseases early [2-3]. The use of sophisticated signal processing algorithms can help extract important features from different physiological signals, such as Electrocardiogram (ECG), Echocardiography, and Magnetic Resonance Imaging (MRI). Advanced neural architectures like convolutional and recurrent models are capable of capturing intricate temporal and spatial features in cardiac signals, thereby enhancing the precision of cardiovascular disease classification.

These techniques have advantages of adopting in heart disease detection, such as improved accuracy, speed and efficiency, scalability, and personalized medicine. The potential impact of developing effective and accurate heart disease detection techniques using signal processing and deep learning techniques is vast. It is instrumental in enhancing the quality and efficiency of healthcare and advancing the field of medical research. An algorithm for detecting premature ventricular contractions based on discrete wavelet transform (DWT) coefficients and probabilistic neural network (PNN) was proposed in [4], which classified eight different heartbeats. [5] presented a diagnostic model that classified four cardiac arrhythmia data classes based on a combination of SVM-based categories and features of wavelet coefficients, local binary patterns, and wave amplitude values. [6] introduced the usage of features of morphology, wavelet transform, and S transform for heart disease diagnosis based on SVM classification and a Gaussian process detector. [7] presented an online algorithm for premature ventricular arrhythmia diagnosis by using the PCA algorithm, which compared the main directions of changes in the heart rate of normal and sick people. [8] evaluated the performance of different routines based on discrete wavelet

ISSN

transform coefficients and PCA analysis for the Recognition of Cardiac Rhythm Irregularities as normal and abnormal heartbeats. [9] utilized 12 morphological coefficients, tiger energy operator coefficients, and features obtained from DWT for neural network training for the diagnosis of heart diseases. Both ML-based and deep learning-based methods have demonstrated strong performance in ECG rhythm classification. The choice of algorithm may depend on the complexity of the ECG signal, the size of the dataset, and the desired accuracy of the classifier. However, deep learning-based methods are more common in today's research and development owing to their superior performance in classification tasks. The advantages of deep learning-based methods include their ability to learn complex and nonlinear interactions between input features, which helps to identify more robust and discriminative features. Additionally, their capacity to handle large amounts of data and extract high-level features has enabled them to achieve state-of-the-art results in many applications, including medical imaging and natural language processing. So, in this paper, an attempt has been made to investigate classifiers based on deep learning methods further.

Recent research has explored a variety of deep learning strategies for automatic ECG interpretation. For instance, [11] introduced a fully automated one-dimensional convolutional network designed to identify five heartbeat categories, incorporating wavelet-based noise reduction as a preprocessing step. In a comprehensive survey, [12] outlined various deep learning frameworks along with data preparation techniques and performance metrics specifically tailored to arrhythmia classification. A study in [13] proposed a hybrid architecture combining two-dimensional CNNs with LSTM layers, where ECG signals were first converted into color-coded scalograms using wavelet transforms to facilitate spatial-temporal learning. In [14], a pre-trained DenseNet model was employed to extract deep representations from ECG-derived spectrograms, which were subsequently classified using a support vector machine (SVM) into four arrhythmia types. Reference [15] utilized a convolutional network to detect seven forms of premature ventricular contractions directly from raw ECG time-series inputs, with classification performed via a softmax decision layer. Lastly, [16] integrated convolutional autoencoders and transfer learning mechanisms to distinguish between normal and pathological signals, relying on their two-dimensional scalogram representations as model inputs.

The proposed method for detecting cardiac arrhythmias involves several steps. The first step is to obtain a 2D encoding of ECG characteristics in both frequency and time domains. A scalogram algorithm is used to break the signal down into its frequency components over time. This representation provides information about the behavior and features of the signal over time. The next step is to extract important frequency-domain features from the scalogram using a 2D CNN. A series of convolutions is used to identify important areas in the signal and extract associated features. The third step is to feed the extracted features from the scalogram to an LSTM network to capture the structure of the ECG signal. The final step is to use a fully connected layer to obtain the final output of the model, indicating the presence or absence of cardiac arrhythmias. The model combines the features extracted from the scalogram and LSTM network and passes them through a fully connected layer. The model ultimately produces a binary value indicating whether cardiac arrhythmias are present in the ECG signal.

Section 2 outlines the methodology adopted for arrhythmia detection using ECG signals, including comprehensive details on the utilized dataset, signal preprocessing pipeline, feature transformation strategies, and configuration of the developed model. The third section provides an in-depth analysis of the experimental procedures along with quantitative performance metrics. Section 4 offers a critical interpretation of the outcomes achieved by the proposed framework. Finally, Section 5 summarizes the key findings and outlines potential directions for future work.

2. The proposed ECG arrhythmia classifier

The developed framework leverages time–frequency characteristics of ECG signals, employing a CNN module for effective feature extraction and an LSTM network for sequential classification. This hybrid strategy integrates signal processing with deep learning to enable accurate identification of cardiac arrhythmias. The method involves several steps, each of which performs a specific task: 1) Obtaining a time-frequency domain representation of the ECG signal: The first step in the process is to obtain a 2D representation of the ECG signal in both frequency and time domains. This is achieved by processing the input ECG signal through a scalogram algorithm, which breaks the signal down into its frequency components over time. This representation provides information about the behavior and features of the signal over time. 2) Extracting features from the scalogram using a 2D CNN: The second step is to extract important frequency-domain features from the scalogram using a 2D CNN. This involves passing the scalogram through a series of convolutions, which helps to identify important areas in the signal and extract associated features. The CNN can learn to localize these features in time, allowing it to identify important patterns and trends in the structure of the signal. 3) Feeding the CNN features to an LSTM network to capture the structure of the ECG signal: The third step involves feeding the extracted features from the scalogram to an LSTM network, which can model the time-frequency structure of the ECG signal. The LSTM network can learn long-term dependencies in the signal and capture important time-frequency features that might not be apparent in the individual frames of the spectrogram. By analyzing these features, the LSTM network can identify patterns and trends that indicate the presence of cardiac arrhythmias. 4) Using a fully connected layer to obtain the final output: The final step is to use a fully connected layer to obtain the final output of the model, indicating the presence or absence of cardiac arrhythmias. This is achieved by combining the features extracted from the scalogram and LSTM network and passing them through a fully connected layer. The model produces a binary output that indicates the presence or absence of cardiac arrhythmias in the ECG signal. Figure 1 illustrates the block diagram of the proposed approach, and the individual components of this diagram are described in detail in the following subsections.

2.1. Dataset

To evaluate the effectiveness of the proposed arrhythmia detection approach, the Physionet arrhythmia datasets were utilized. These datasets contain labeled cardiac signals corresponding to different types of arrhythmias [17]. Physionet provides access to diverse ECG datasets that can be used to train and evaluate deep learning models for cardiovascular disease detection, diagnosis, and monitoring. The datasets typically include information about the patients, such as their medical history and demographic information, which can be used to label and categorize the ECG signals. The dataset includes 22 heart recordings from female subjects aged between 23 and 89 years, as well as 25 recordings from male subjects aged 22 to 79 years [17]. From the MIT-BIH cardiac arrhythmia database [18], a total of 96 recordings were selected, featuring sampling rates of 128 and 360 Hz. Additionally, the MIT-BIH Normal Sinus Rhythm database provides 30 recordings sampled at 128 Hz with 12-bit resolution. A typical normal heartbeat waveform, which consists of several distinct wave components, is illustrated in Figure 2.

2.2. Pre-processing steps

In the ECG signal preprocessing phase, various operations are applied to condition the data for subsequent training of statistical or deep learning models. Since ECG recordings often contain diverse types of noise and artifacts, such as interference from power lines, muscle activity, power-line interference (PLI), and electromyographic (EMG) signals, it is essential to eliminate these disturbances. Both low-frequency and high-frequency noise components need to be suppressed before proceeding to feature extraction or model training. To achieve this, multiple denoising techniques have been proposed, including classical filter-based approaches, wavelet transform methods, and hybrid algorithms that integrate the outputs of multiple filters to enhance noise removal effectiveness [19–21].

In this study, noise removal is achieved by first centering the cardiac signals, which involves subtracting the mean value from the ECG data. Additionally, an adaptive median filter is applied to smooth the raw signals. To extract the relevant features, the detection of heartbeats is necessary. Deep learning techniques demand extensive training data, and in the Physionet arrhythmia database, each heartbeat is annotated with a specific arrhythmia label. For classification purposes, this work focuses on three types of ECG rhythms: cardiac arrhythmia (ARR), congestive heart failure (CHF), and normal sinus rhythm (NSR). Each ECG recording is segmented into 360 samples centered on the R-peak of each heartbeat, as described previously.

2.3. Data conversion and mapping based on CWT

The raw time ECG signal usually contains many types of noise and artifacts that can make the CNN-based classification more challenging and less reliable.

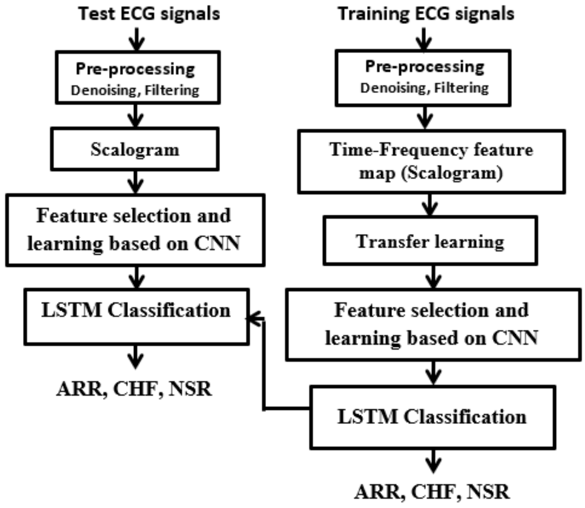


Figure 1. Schematic overview of the ECG classification framework combining CNN and LSTM models.

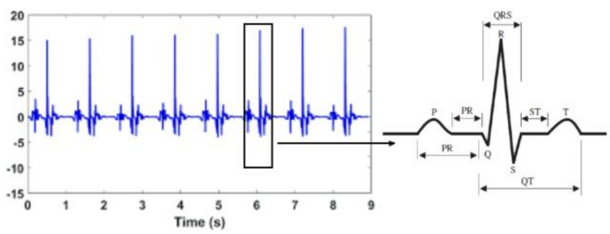


Figure 3. A normal heartbeat signal consists of several different wave components.

One of the main reasons why a raw time ECG signal is not properly classified using a CNN is that CNNs are designed to process images, and a raw ECG signal is a time-series signal. In order to provide the CNN with appropriate input data, the ECG signals are converted into the time-frequency domain. This transformation enables the CNN to efficiently capture and learn features relevant to the distinct patterns within the ECG signals.

The transformation helps to reduce the impact of noise and artifacts and also improves the performance of the CNN-based classifier. There are several ways to transform an ECG signal into a time-frequency domain suitable for 2D CNN classification. One popular way is by applying CWT, which splits the ECG signal into several subbands representing different frequencies and time scales as scalograms. A scalogram visually depicts how the frequency content of a signal evolves by representing it simultaneously in the time and frequency domains. The scalogram is created by applying a CWT to the signal, which analyzes the signal over different scales and frequencies. A CWT analysis of an ECG signal can provide valuable information about the rhythm of the heart and the underlying cardiac activity. The selection of wavelet type for a CWT analysis depends on the desired level of resolution in both the time and frequency domains. Different wavelet types can be used to highlight different features of the ECG signal. Common wavelet types used for ECG signal processing include the Haar, Daubechies, Morlet, and Symmlet wavelets. In this paper, the Daubechies wavelet was employed to provide a good time-frequency resolution and suppression of noise. Each subband contains a representation of the ECG signal that includes information about both its spatial (time) and spectral (frequency) aspects. The CNN operates on small windows of the time-frequency content, extracting important features and localizing them. Figure 3 demonstrates the conversion of raw ECG recordings from various subjects, belonging to three categories, cardiac arrhythmia (ARR), congestive heart failure (CHF), and normal sinus rhythm (NSR), into scalograms that represent the signals in the time-frequency domain. This conversion is performed through the wavelet transform, an effective mathematical technique that breaks down the signal into frequency components across multiple time scales. This transformation provides a more visually appealing representation of the ECG signal, with each point in the scalogram representing the signal's power or frequency content in a specific time-scale region. The Figures show that the transformed signals and scalograms for each of the three classes are visually distinct, with different patterns and frequency content, which can be effectively utilized to train and classify the CNN model.

2.4. Data augmentation

In CNN-based ECG arrhythmia classification, data augmentation plays a crucial role by expanding the training dataset both in size and variety, thereby enhancing the model's ability to generalize and improving overall classification accuracy. The limited amount of available ECG data is one of the main challenges in ECG-based classification, especially in cases of rare arrhythmias. Data augmentation techniques can generate new data by applying various transforms, such as translation, rotations, scaling, shearing, and centering, to the available ECG data [22-23]. By training on an expanded and varied dataset, these augmented samples help strengthen the CNN model's resilience, reducing the risk of overfitting and enhancing its ability to perform well on previously unseen data. Additionally, some augmentation techniques can also help to simulate the effects of real-world ECG signals, such as noise and artifacts, which makes the model more robust to such challenges in real-world scenarios. Table 1 reports the number of ECG frames included in the dataset before and after data augmentation.

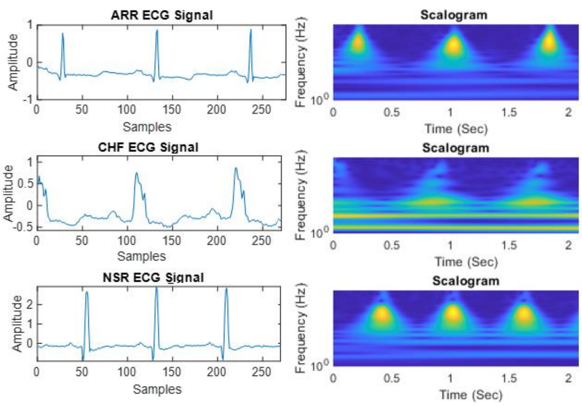


Figure 3. Left panel: Samples of ECG waveforms representing three arrhythmia categories, ARR, CHF, and NSR. Right panel: Corresponding scalogram images of these ECG signals generated using the CWT.

Table 1. Counts of ECG recordings obtained from the Physionet dataset, shown separately for original and augmented data during both training and testing phases.

Classes of ECG signals	Total	After data augmentation	# of Train	# of Test
ARR	15817	22512	15759	6753
CHF	14312	21136	14795	6341
NSR	12523	20322	14225	6097
Total	42652	63970	44779	19191

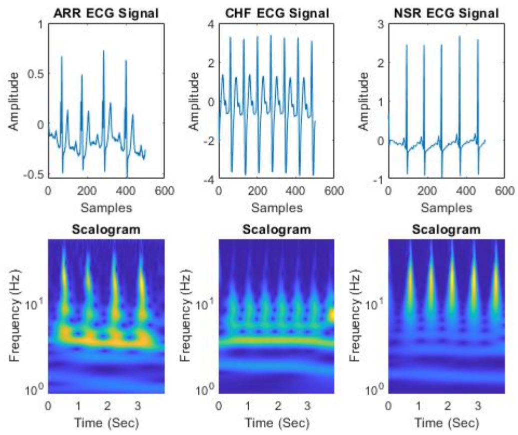


Figure 4. Illustrative examples of ECG data augmentation applied to the ARR, CHF, and NSR classes to enhance dataset variability.

Also, Figure 4 shows augmented ARR, CHF, and NSR signals from Physionet arrhythmia datasets using a scaling factor.

2.5. Architecture of the deep model

In this Section, the proposed algorithm is described as an AI-powered system that employs CNN to classify arrhythmia types of CVDs, namely NSR, CHF, and ARR. CNN is a deep learning architecture that is commonly used in ECG classification because of its ability to learn hierarchical features from raw signal data. CNNs consist of stacked convolutional and pooling layers that progressively extract and integrate features from input signals. Meanwhile, LSTM networks, a specialized form of RNNs, are tailored to process sequential data like ECG signals or speech [24]. It was developed to address the problem of vanishing gradients in RNNs, which leads to poor performance on long-term dependencies in the data. A distinctive characteristic of LSTM networks is their internal structure, composed of three gates: the input gate, forget gate, and output gate. These gates regulate the flow of information by deciding what to retain, discard, or pass on at every time step. This allows LSTM to remember information for an extended period and learn temporal patterns in the data. Unlike traditional RNNs, LSTM networks excel at capturing long-range dependencies within data, which is especially valuable for applications like ECG signal analysis, where understanding the temporal patterns is crucial for precise classification. In this paper, CNN and LSTM were employed to extract the best feature vectors and detect the label of the input ECG signal arrhythmia, respectively. Figure 5 demonstrates the detailed architecture of an LSTM network.

At each time step, the network processes the current input sequence Xt along with the previous hidden state ht-1. A sigmoid activation function, represented by σ, is applied. The forget gate vector ft decides how much information from earlier inputs should be retained, while Ct denotes the prior cell state connected to the current cell state Ct, which corresponds to the input at time t. The network processes sequential data, making it suitable for various tasks, such as video classification and sentiment analysis, that involve complex patterns [24]. Table 2 showcases a fine-tuned LSTM setup and highlights key parameters set for the optimal performance of the proposed method. This Table provides information to evaluate and adjust the parameters to achieve the desired performance outcomes. A stochastic gradient descent algorithm (SGD) known as an adaptive moment estimation (Adam) optimizer with an initial learning rate of 0.001 is employed to train our CNN model. At epochs 20 and 40, the optimizer is multiplied by 0.1, which is a strategy to adaptively adjust the learning rate during training [25]. To reduce the risk of overfitting, the learning rate is halved every 10 epochs. Training is conducted using mini-batch gradient descent with a batch size of 64 samples per iteration. The network undergoes training for up to 100 epochs or until the loss function reaches convergence.

Table 3 summarizes the network hyperparameters for the described CNN model in each scenario. Transfer learning is a technique in machine learning that involves using knowledge gained from a pre-trained model on a similar task to help a new model learn to perform a related task. In the context of ECG arrhythmia classification, transfer learning can be used to train a deep model using a pre-existing ECG dataset, which helps the new model to learn faster and improve its performance.

3. Simulation results of the proposed method

This study presents an intelligent framework for classifying ECG arrhythmias by transforming raw signals, employing a deep learning model, and benchmarking against advanced techniques. Initially, the denoised and segmented ECG time series were converted into two-dimensional scalogram images. These scalograms served as input to a CNN model, designed with an input layer dimension of 256×256×3. The classification result was derived from the output of a fully connected layer following batch normalization.

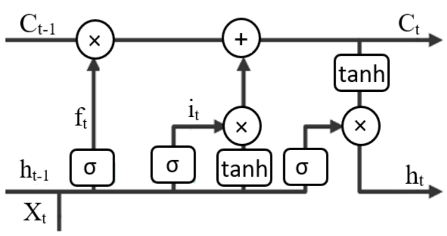


Figure 5. The framework of a simple LSTM network.

Table 2. The fine-tuned parameters of the employed LSTM deep model in the proposed ECG classifier algorithm.

Parameter	Description	Range
Number of LSTM layers	Number of LSTM units stacked vertically to control long-term memory capacity	2
Number of units per layer	Number of LSTM cells in each layer to controls model complexity and expressiveness	64
Input dimensionality	Size of the input vector for data representation	3×256×256
Activation function	Non-linear transformation applied to activation	Sigmoid
Optimizer	An algorithm for adjusting model weights to determine learning rate and convergence behavior	Adam
Loss function	The metric used to evaluate model performance	Mean squared error
Regularization techniques	Methods to prevent overfitting and control model complexity and generalization	Dropout
Learning rate	Controls how much weights are updated and affect convergence speed and stability	0.001
Batch size	Number of samples processed together to balances efficiency and optimization stability	64

Table 3. Optimal hyperparameters for the employed deep models.

Hyperparameters	CNN	CNN-LSTM	Transferred CNN-LSTM
Batch size	32	64	64
Epochs	50	100	100
Optimizer	SGD	ADAM	ADAM
Learning rate	0.01	0.001	0.001
Interpolate method	Linear	N/A	N/A
Kernel size	3×3	3×3	3×3
Activation function	Sigmoid	Sigmoid	Sigmoid
Weight decay	10^{-3}	10^{-4}	10^{-4}
Dropout	50%	25%	25%

[27] proposed a technique that combines specific morphological features with coefficients obtained from various decomposition levels of the wavelet packet transform to achieve data dimensionality reduction. Subsequently, the reduced data were modeled using the Semi-Nonnegative Matrix Factorization (SNMF) algorithm. The study evaluated model performance using several metrics, including Accuracy (Acc), defined as the proportion of correctly classified instances relative to the total number of samples:

$$\mathrm{Acc} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}) \tag{1}$$

In this context, TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives, respectively. True positives correspond to correctly identified arrhythmia instances, while true negatives are non-arrhythmia cases accurately recognized as such. False positives occur when normal instances are mistakenly classified as arrhythmias, and false negatives refer to arrhythmia cases that the model fails to detect, labeling them as normal instead. 2) Specificity (Spe) stands for the percentage of true negatives in all negatives. 3) Precision (Ppr) for the percentage of true positives in all positive classifications. 4) Sensitivity (Sen or R) calculates the percentage of true positives in all cases. Also, the F-measure that measures combines accuracy, sensitivity, and precision to give a single performance metric as

$$\mathrm{Spe} = \mathrm{TN}/(\mathrm{TN}+\mathrm{FP}) \quad (2)$$
$$\mathrm{Ppr} = \mathrm{TP}/(\mathrm{TP}+\mathrm{FP}) \quad (3)$$
$$\mathrm{Sen}\ (\mathrm{R}) = \mathrm{TP}/(\mathrm{TP}+\mathrm{FN}) \quad (4)$$
$$\mathrm{F}-\mathrm{measure} = 2 \times (\mathrm{Ppr}\ \mathrm{Sen})/(\mathrm{Ppr}+\mathrm{Sen}) \quad (5)$$

The F-measure represents a harmonic mean of precision and sensitivity, metrics that hold particular significance in medical contexts where minimizing false positives is often more critical than false negatives. Higher values of these indicators correspond to improved classification performance.

The proposed algorithm for learning the deep model is trained using 5-fold cross-validation, and overall accuracy, positive prediction rate, sensitivity, and F-measure, confusion matrix are referenced in Table 6. Also, a confusion matrix is a useful statistical tool that gives an overview of the performance of a classifier model based on its ability to classify datasets into different classes. This matrix can help to identify the strengths and weaknesses of a classifier model and to make improvements where necessary. The confusion matrix of the proposed classifier with a deep model designed in Scenario III is reported in Table 5. Also, to accurately evaluate the performance of the proposed algorithm, Friedman's statistics are used to compare the efficiency of different methods. The null hypothesis assumes that all the compared methods perform equally effectively. A significance level (p-value) of 0.05 was chosen to assess whether observed differences among the methods are statistically meaningful. The ρ-value represents the probability of obtaining results as extreme as those observed, assuming the null hypothesis of no difference holds. As shown in Table 7, all reported ρ-values fall below the 0.05 threshold, providing strong evidence to reject the null hypothesis. This indicates that the proposed algorithm outperforms the other approaches listed in Table 7 in classifying ECG arrhythmias. Moreover, the CNN-LSTM-based model demonstrated not only high accuracy but also consistent performance across metrics such as specificity, positive predictive value, sensitivity, and F-measure on the evaluated datasets. Testing on the Physionet ECG database confirmed state-of-the-art results, with the model achieving an accuracy of 98.71%, sensitivity of 98.93%, and a positive prediction rate of 98.99%. The results show that the proposed approach is effective in accurately detecting ECG heart abnormalities. Furthermore, the proposed method is highly efficient and can process large ECG data in real time with high accuracy. The main advantage of the proposed approach is that it combines multiple deep learning techniques such as CNN, LSTM, and TL to achieve better results in ECG classification. Also, the ability to use TL helps to improve the model's performance and generalize it to new environments without losing accuracy. The proposed approach can be further improved by exploring new algorithms, increasing the size of the datasets, and fine-tuning the network parameters.

Table 4. Comparison of classification accuracy (Acc) for ECG arrhythmia across various scenarios using the proposed approach versus alternative techniques.

	ARR	CHF	NSR	Mean
RNN-based [13]	96.70	97.82	97.39	97.30
MLP-based [26]	97.33	97.21	97.72	97.42
SNMF-based [27]	95.51	95.38	96.41	95.77
CNN (Scenario I)	97.99	98.05	97.93	97.99
CNN-LSTM (Scenario II)	98.56	98.49	98.62	98.56
CNN-LSTM-TL(Scenario III)	**98.81**	**98.60**	**98.86**	**98.76**

Table 5. Summary of ECG arrhythmia classification performance metrics, including Accuracy (Acc), Positive Prediction Rate (Ppr), Sensitivity (Sen), and F-measure, achieved by the proposed RenNet34-based method.

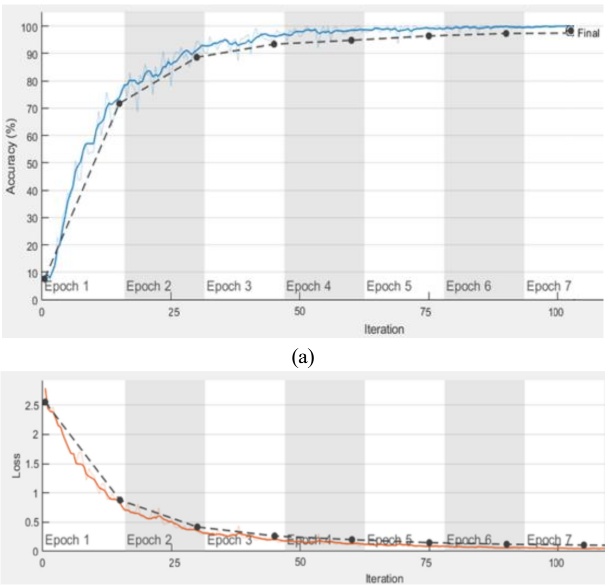
		Acc (%)	Ppr (%)	Sen (%)	F-Measure (%)
RNN-based [13]	ARR	96.70	97.10	97.32	97.20
	CHF	97.82	97.67	96.77	97.21
	NSR	97.39	96.75	97.65	97.20
MLP-based [26]	ARR	97.33	97.29	97.04	97.16
	CHF	97.21	97.50	97.97	97.73
	NSR	97.72	96.71	97.15	96.93
SNMF-based [27]	ARR	95.51	96.61	96.89	96.74
	CHF	95.38	96.37	96.44	96.40
	NSR	96.41	96.58	96.51	96.54
CNN (Scenario I)	ARR	97.99	98.13	98.08	98.10
	CHF	98.05	98.10	98.79	98.44
	NSR	98.93	96.57	98.27	97.41
CNN-LSTM (Scenario II)	ARR	98.56	97.53	98.97	98.24
	CHF	98.49	98.48	98.69	98.58
	NSR	98.62	98.59	98.56	98.57
CNN-LSTM-TL (Scenario III)	ARR	**98.67**	**98.89**	**98.72**	**98.80**
	CHF	**98.60**	**98.96**	**99.00**	**98.98**
	NSR	**98.86**	**99.11**	**99.06**	**99.08**

Table 6. Confusion matrices illustrating the classification results for ECG arrhythmia across various CNN-based scenarios.

ARR(class 1)	**98.67%**	0.53%	0.80%
CHF(class 2)	0.69%	**98.60%**	0.71%
NSR(class 3)	0.31%	0.83%	**98.86%**
	ARR(class 1)	CHF(class 2)	NSR(class 3)

Table 7. **Results of statistical analyses comparing different methods alongside the CNN-based scenarios.**

		ρ-value	Mean ρ-value	Mean Acc (%)
RNN-based [13]	ARR	0.004		
	CHF	0.005	0.0046	97.30
	NSR	0.005		
MLP-based [26]	ARR	0.007		
	CHF	0.006	0.0060	97.42
	NSR	0.005		
SNMF-based [27]	ARR	0.011		
	CHF	0.008	0.0096	95.77
	NSR	0.010		
CNN (Scenario I)	ARR	0.004		
	CHF	0.003	0.0036	98.32
	NSR	0.004		
CNN-LSTM (Scenario II)	ARR	0.002		
	CHF	0.003	0.0023	98.56
	NSR	0.002		
CNN-LSTM-TL (Scenario III)	ARR	**0.002**		
	CHF	**0.001**	**0.0013**	**98.71**
	NSR	**0.001**		



(a)

(b)

Figure 6. **Training curves for the proposed CNN-LSTM model in Scenario III: a) Accuracy progression during training, b) Loss function trend illustrating model fitting improvement. Dashed lines correspond to validation dataset metrics.**

In summary, the proposed method is an effective combination of CNN, LSTM, and TL that achieves high accuracy in detecting ECG heart abnormalities. The proposed approach is highly efficient and real-time, and can process large ECG data accurately. Furthermore, its ability to generalize to new environments without sacrificing accuracy makes it an attractive option for healthcare applications. Figure 6 illustrates the training progress of the proposed CNN-LSTM model integrated with transfer learning in the third scenario. These plots provide insight into the model's learning dynamics and assist in tuning hyperparameters for improved classification performance. The results indicate that the model achieves high accuracy in distinguishing various arrhythmia types, underscoring its potential utility in medical diagnostics.

4. Conclusion

This research introduces a reliable deep learning-based approach for detecting cardiac arrhythmias aimed at facilitating early diagnosis of heart conditions. Given the rising prevalence and serious consequences of cardiovascular diseases, timely identification is essential for effective treatment and prevention. The proposed framework integrates CNN and LSTM networks to analyze ECG signals in the time-frequency domain, utilizing CWT. Initially, the ECG signals are converted into scalogram images via CWT, providing a two-dimensional representation that captures both temporal and spectral information. These scalograms serve as inputs to the CNN, which extracts spatial features, while the LSTM component models the sequential dependencies inherent in heartbeat patterns. The method achieved impressive accuracy levels, reaching 98.76%, indicating its promise as a diagnostic tool. By enhancing early detection capabilities, this approach can support clinicians in improving patient outcomes and potentially lowering healthcare costs related to heart disease management. This technique marks a notable advancement over conventional signal processing methods for arrhythmia detection, combining effective feature extraction with temporal modeling to boost both accuracy and efficiency.

References

[1] Xiao, Q., Lee, K., Mokhtar, S. A., Ismail, I., Pauzi, A. L. M., Zhang, Q., & Lim, P. Y. (2023). Deep learning-based ECG arrhythmia classification: A systematic review. *Applied Sciences, 13*(8), 4964. https://doi.org/10.3390/app13084964.

[2] Ebrahimi, Z., Loni, M., Daneshtalab, M., & Gharehbaghi, A. (2020). A review on deep learning methods for ECG arrhythmia classification. *Expert Systems with Applications, 7*.

[3] Hong, S., Zhou, Y., Shang, J., Xiao, C., & Sun, J. (2020). Opportunities and challenges of deep learning methods for electrocardiogram data: A systematic review. *Computers in Biology and Medicine, 122*.

[4] Gutiérrez-Gnecchi, J. A., Morfin-Magaña, R., Lorias-Espinoza, D., et al. (2017). DSP-based arrhythmia classification using wavelet transform and probabilistic neural network. *Biomedical Signal Processing and Control, 32*, 44–56.

[5] Mondéjar-Guerra, V., Novo, J., Rouco, J., Penedo, M. J., & Ortega, M. (2019). Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers. *Biomedical Signal Processing and Control, 47*, 41–48.

[6] Alajlan, N., Bazi, Y., Melgani, F., Malek, S., & Bencherif, M. A. (2014). Detection of premature ventricular contraction arrhythmias in electrocardiogram signals with kernel methods. *Signal, Image and Video Processing, 8*(5), 931–942.

[7] Zarei, R., He, J., Huang, G., & Zhang, Y. (2016). Effective and efficient detection of premature ventricular contractions

based on variation of principal directions. *Digital Signal Processing, 50*, 93–102.

[8] Kaur, I., Rajni, R., & Marwaha, A. (2016). ECG signal analysis and arrhythmia detection using wavelet transform. *Journal of the Institution of Engineers (India): Series B, 97*(4), 499–507.

[9] Anwar, S. M., Gul, M., Majid, M., & Alnowami, M. R. (2018). Arrhythmia classification of ECG signals using hybrid features. *Conference Paper*, 1–8.

[10] Jekova, I., & Krasteva, V. (2020). Optimization of end-to-end convolutional neural networks for analysis of out-of-hospital cardiac arrest rhythms during cardiopulmonary resuscitation. *Sensors, 21*.

[11] Wu, M., Lu, Y., Yang, W., & Wong, S. Y. (2021). A study on arrhythmia via ECG signal classification using the convolutional neural network. *Frontiers in Computational Neuroscience, 14*.

[12] Xiao, Q., Lee, K., Mokhtar, S. A., Ismail, I., Pauzi, A. L. M., Zhang, Q., & Lim, P. Y. (2023). Deep learning-based ECG arrhythmia classification: A systematic review. *Applied Sciences, 13*(8), 4964.

[13] Madan, P., Singh, V., Singh, D. P., Diwakar, M., Pant, B., & Kisho, A. (2022). A hybrid deep learning approach for ECG-based arrhythmia classification. *Bioengineering, 9*(4).

[14] Salem, M., Taheri, S., & Yuan, J. S. (2018). ECG arrhythmia classification using transfer learning from 2-dimensional deep CNN features. In *IEEE Biomedical Circuits and Systems Conference (BioCAS)*, 1–4.

[15] Rajkumar, A., Ganesan, M., & Lavanya, R. (2019). Arrhythmia classification on ECG using deep learning. In *5th International Conference on Advanced Computing & Communication Systems (ICACCS)*, IEEE, 365–369.

[16] Obaidi, R. M., Sattar, A. R., Abd, M., Almani, I., Alghazali, T., Talib, S., Ghazi, A., Mohammad, M. Q., Abid, T., & Abdul Sahib, M. R. (2022). ECG arrhythmia classification based on convolutional autoencoders and transfer learning. *MAJLESI Journal of Electrical Engineering, 16*(3), 41–46.

[17] PhysioNet Database Archive. Available: https://archive.physionet.org/physiobank/database.

[18] The MIT-BIH Arrhythmia Database. (2015, October 8). Available: http://physinet.org/physiobank/database/mitdb.

[19] Oh, S., & Lee, M. (2022). A shallow domain knowledge injection (SDK-Injection) method for improving CNN-based ECG pattern classification. *Applied Sciences, 12*.

[20] Alkhodari, M., Apostolidis, G., Zisou, C., Hadjileontiadis, L. J., & Khandoker, A. H. (2021). Swarm decomposition enhances the discrimination of cardiac arrhythmias in varied-lead ECG using ResNet-BiLSTM network activations. In *Proceedings of the 2021 Computing in Cardiology (CinC)*, Brno, Czech Republic, 13–15.

[21] Liu, Z., Zhou, B., Jiang, Z., Chen, X., Li, Y., Tang, M., & Miao, F. (2022). Multiclass arrhythmia detection and classification from photoplethysmography signals using a deep convolutional neural network. *Journal of the American Heart Association, 11*.

[22] Rahman, M. M., Rivolta, M. W., Badilini, F., & Sassi, R. (2023). A systematic survey of data augmentation of ECG signals for AI applications. *Sensors, 23*. https://doi.org/10.3390/s23115237.

[23] Ma, S., Cui, J., Xiao, W., & Liu, L. (2022). Deep learning-based data augmentation and model fusion for automatic arrhythmia identification and classification algorithms. *Computational Intelligence and Neuroscience*. https://doi.org/10.1155/2022/1577778.

[24] Tatsunami, Y., & Taki, M. (2022). Sequencer: Deep LSTM for image classification. *arXiv preprint arXiv:2205.01972*.

[25] Kingma, D., & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.

[26] Savalia, S., & Emamian, V. (2018). Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks. *Bioengineering (Basel), 5*(2). https://doi.org/10.3390/bioengineering5020035.

[27] Mavaddati, S. (2020). ECG arrhythmia classification based on wavelet packet transform and sparse non-negative matrix factorization. *Journal of Iranian Association of Electrical and Electronics Engineers, 17*(3), 119-128.