

Article

Enhanced Classification of Heartbeat Electrocardiogram Signals Using a Long Short-Term Memory–Convolutional Neural Network Ensemble: Paving the Way for Preventive Healthcare

Njud S. Alharbi ¹ , Hadi Jahanshahi ^{2,*} , Qijia Yao ³ , Stelios Bekiros ^{4,5,6}  and Irene Moroz ⁷

¹ Department of Biological Sciences, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia; nsfalharbi@kau.edu.sa

² Institute of Electrical and Electronics Engineers, Toronto, ON M5V 3T9, Canada

³ School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China; qijia_yao@ustb.edu.cn

⁴ Department of Banking and Finance, FEMA, University of Malta, MSD 2080 Msida, Malta; stelios.bekiros@um.edu.mt

⁵ LSE Health, Department of Health Policy, London School of Economics and Political Science, London WC2A 2AE, UK

⁶ IPAG Business School, 184, Bd Saint-Germain, 75006 Paris, France

⁷ Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK; irene.moroz@maths.ox.ac.uk

* Correspondence: jahanshahi.hadi@ieee.org

Abstract: In the rapidly evolving field of medical diagnosis, the accurate and prompt interpretation of heartbeat electrocardiogram (ECG) signals have become increasingly crucial. Despite the presence of recent advances, there is an exigent need to enhance the accuracy of existing methodologies, especially given the profound implications such interpretations can have on patient prognosis. To this end, we introduce a novel ensemble comprising Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models to enable the enhanced classification of heartbeat ECG signals. Our approach capitalizes on LSTM's exceptional sequential data learning capability and CNN's intricate pattern recognition strength. Advanced signal processing methods are integrated to enhance the quality of raw ECG signals before feeding them into the deep learning model. Experimental evaluations on benchmark ECG datasets demonstrate that our proposed ensemble model surpasses other state-of-the-art deep learning models. It achieves a sensitivity of 94.52%, a specificity of 96.42%, and an accuracy of 95.45%, highlighting its superior performance metrics. This study introduces a promising tool for bolstering cardiovascular disease diagnosis, showcasing the potential of such techniques to advance preventive healthcare.

Keywords: cardiovascular disease diagnosis; ensemble neural network; time series classification; convolutional neural network; recurrent neural network

MSC: 68Txx; 92Cxx; 65Cxx; 68Uxx; 92Bxx



Citation: Alharbi, N.S.; Jahanshahi, H.; Yao, Q.; Bekiros, S.; Moroz, I. Enhanced Classification of Heartbeat Electrocardiogram Signals Using a Long Short-Term Memory–Convolutional Neural Network Ensemble: Paving the Way for Preventive Healthcare. *Mathematics* **2023**, *11*, 3942. <https://doi.org/10.3390/math11183942>

Academic Editors: Jérôme Mendes, Alireza Mousavi, Pedro Manuel Soares Moura, António J. Baptista and Morad Danishvar

Received: 4 August 2023

Revised: 6 September 2023

Accepted: 15 September 2023

Published: 17 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Heart disease has emerged as one of the most pressing health challenges both globally and nationally [1,2]. Cardiovascular ailments account for a significant proportion of mortality rates, emphasizing the persistent and widespread nature of this medical concern. In some urban populations, changing lifestyles and genetic predispositions have contributed to a surge in cardiac-related issues, making them a primary public health focus [3,4]. As these diseases continue to be a major concern, the importance of technological advancements in diagnostic tools becomes even more evident. ECG signals, which depict the electrical activity of the heart over a period of time, have long been a cornerstone in the clinical detection of heart diseases [5,6]. However, the interpretation of these ECG signals is not

straightforward and requires extensive medical knowledge and experience. Furthermore, in our current age of digital healthcare, vast amounts of ECG data are generated, necessitating the use of automated techniques that can efficiently process and interpret these signals to aid clinicians in their diagnostic processes. In contemporary medical practice, machine learning has emerged as an indispensable tool in the diagnosis of diseases [7,8]. It demonstrates the ability to process and interpret voluminous medical data swiftly, thereby enhancing diagnostic accuracy and efficiency. It finds its application in the analysis of intricate medical images—CT scans, MRIs, and X-rays—to identify minute patterns associated with diseases [9]. These capabilities make the early detection of conditions such as cancer, neurodegenerative diseases, and cardiac ailments feasible [10].

While ECG is a pivotal tool across various fields, it confronts certain inherent challenges that limit its in-depth analysis. For example, the ECG signal's non-stationary nature [11] indicates that its statistical features can change over time. Consequently, models that are trained using data from a specific time frame might struggle to effectively interpret data from a different time period, even if sourced from the same individual. This dynamic quality of ECG poses considerable obstacles to real-world applications. Furthermore, there is significant variability in ECG data between individuals. These variations, linked to diverse patterns among subjects [12], can severely compromise the performance of models designed for a wider audience.

The use of machine learning in genomics enables the recognition of disease-associated genetic patterns and mutations, thereby facilitating personalized treatment approaches [13,14]. A particularly noteworthy application is predictive analytics, which uses machine learning to forecast disease outbreaks and progressions based on historical and real-time data, thus contributing significantly to public health management [15]. However, traditional machine learning methods, despite their significant advancements, have demonstrated certain limitations when applied to ECG signal classification. Firstly, traditional machine learning techniques often require manual feature extraction from ECG signals, which can be time-consuming and may not capture all the nuanced information present in these signals [16]. Secondly, ECG signals can have a vast range of morphologies due to patient-to-patient variability, and traditional methods might not always generalize well across diverse datasets. Thirdly, ECG signals are inherently non-linear and non-stationary, and traditional linear models can struggle to capture these dynamics effectively. Lastly, the rapid evolution of wearable devices producing continuous ECG monitoring necessitates models that can process large-scale data efficiently, adapt in real-time, and make instant predictions—tasks that might be beyond the scope of classical machine learning algorithms. These challenges have catalyzed the need for more advanced methods, such as deep learning architectures, which can automatically extract features, model complex relationships, and scale more effectively with large datasets [17,18]. These methods are, however, often sensitive to noise and lack robustness, leading to inconsistency in detecting complex patterns in ECG signals [19,20]. Furthermore, these algorithms may struggle to capture the intricate temporal dependencies inherent in ECG signals, often leading to significant information loss and, consequently, reduced diagnostic accuracy.

The complexities of these issues have necessitated ongoing and vibrant research efforts, capitalizing on innovative ideas, methodologies, and advancements. For instance, in recent years, deep learning models have demonstrated exceptional capabilities in an array of sequence and pattern recognition tasks. Their proficiency in these areas positions them as particularly promising tools for the interpretation of ECG signals. Specifically, LSTM [21] models have proven adept at handling sequential data and capturing long-term dependencies in temporal sequences. On the other hand, CNN [22] models have shown superior performance in detecting local patterns and learning spatial hierarchies directly from complex data. While there have been advancements and certain models excel in their respective areas, applying them independently to ECG signal classification might result in overlooking essential features that could be captured by the other. Furthermore, the

accuracy of many methods is not yet as robust as required, limiting their trustworthiness for unsupervised real-world applications.

Considering these observations, the motivation for this study is rooted deeply in the aspiration to overcome the prevalent limitations of current methodologies. By recognizing that a singular approach often results in missed opportunities for comprehensive signal analysis, we propose an innovative ensemble of LSTM and CNN models tailored for the nuanced classification of ECG signals. This integrated framework not only amalgamates the strengths of both LSTM and CNN models, but also ensures a holistic capture of both temporal dynamics and spatial intricacies inherent in ECG signals. The end goal is a substantial improvement in the accuracy of heartbeat classification, setting a new standard in the field. Beyond the amalgamation of LSTM and CNN models, the study also places a significant emphasis on the quality of input data. Recognizing that the fidelity and reliability of the classification largely hinge on the quality of input signals, we employ advanced signal processing techniques to refine and enhance ECG signal inputs, ensuring that the deep learning model receives data of the highest possible quality. Through these concerted efforts, our study endeavors to pave the way for new research in ECG signal classification, where accuracy and reliability are paramount.

In this paper, we present a unique ensemble classification technique for ECG signals that offers several notable contributions. First, our method introduces a new ensemble framework expressly designed for time series classification, which integrates LSTM, bidirectional LSTM (BiLSTM), and CNN models. This fusion capitalizes on the unique strengths of each model to enhance classification accuracy. Second, our methodology exhibits efficiency with its ability to achieve rapid processing speeds, thus facilitating real-time or near-real-time applications. Third, our ensemble model counteracts the issue of overfitting by utilizing the diversity among individual models, leading to improved robustness and generalization. Furthermore, the method we present is resilient against noise, demonstrating its applicability in real-world ECG signal classification scenarios. In this paper, we provide a detailed description of our approach and present the outcomes of experiments performed on benchmark ECG datasets. These results highlight the superior performance of our proposed ensemble model compared to traditional machine learning and standalone deep learning models in terms of accuracy, sensitivity, and specificity. By introducing this novel approach, we aspire to contribute significantly to the prompt and effective diagnosis of heart diseases.

2. Neural Networks for Classification

Due to their inherent ability to process data sequentially, Recurrent Neural Networks (RNNs) have garnered recognition as a formidable asset for classification endeavors, specifically when the importance lies in the time-oriented dynamics of the input [23,24]. Different from conventional feedforward networks, the design of RNNs incorporates feedback loops, enabling data to loop within the network over time and endowing it with the ability to retain previous inputs [25,26]. This particular trait enables RNNs to utilize temporal dependencies, which makes them highly proficient at tasks like speech recognition, sentiment analysis, and forecasting time series. By teaching an RNN with suitably labeled data, it can categorize unfamiliar sequences based on patterns it has learned, demonstrating its application in a wide range of classification scenarios.

2.1. LSTM

Traditional RNNs face challenges in capturing long-term dependencies due to the vanishing gradient problem, making them less effective for longer sequences [27,28]. LSTM was introduced by Hochreiter and Schmidhuber in 1997 [29] as a solution to this problem. In contrast to conventional RNNs, LSTM networks are specifically engineered to capture proficiently both long-range dependencies and temporal changes in sequential data [30]. Its unique architecture, which includes memory cells, input gates, output gates, and forget gates, allows it to maintain information over longer sequences without significant decay,

making it suitable for tasks that require a memory of past events [29]. The inherent gating systems within LSTM permit the selective preservation and application of pertinent contextual data, making it particularly well suited to model intricate patterns within time series data [31].

LSTM networks hold several key strengths when it comes to time series classification tasks, as outlined in [32]. One of the primary advantages is their ability to manage long-term dependencies, making them an excellent tool for identifying extended temporal patterns accurately. This is particularly useful in situations where the historical context of the data is critical. Furthermore, LSTM networks have the unique ability to learn and store information over time, facilitating their capacity to model and adjust to evolving patterns in the time series data. This adaptability becomes critical in fluid environments where data characteristics can shift. Lastly, one of the noteworthy benefits of LSTM networks is their capability to automatically identify relevant features from the input time series. This feature minimizes the necessity for manual feature engineering, thus reducing the possibility of overlooking vital information [33]. In what follows, we delve into some of the most recent and pivotal studies on LSTM.

Fischer et al. [34] and Ouyang [35] employed LSTM models to forecast financial markets. Financial data, characterized by intricate features such as non-linearity, non-stationarity, and sequence correlation, pose considerable challenges for forecasting. Fischer [34] ascertained that the LSTM network stands superior to traditional benchmarks, including random forest, standard deep neural networks, and logistic regression. This assertion of LSTM's supremacy is echoed in petroleum production predictions by Sagheer and Kotb [36]. Their innovative approach of stacking multiple LSTM blocks hierarchically enhanced the model's prowess in processing temporal tasks and in grasping data sequence structures. Similarly, while aiming to predict oil market prices, Cen and Wang [37] leveraged the vanilla LSTM architecture and confirmed its superiority in forecasting. Liu [38] extended this exploration by contrasting LSTM with models like support vector machines, focusing on estimating financial stock volatility. The findings spotlighted LSTM's relatively straightforward calibration process and its ability to accurately predict over extended time intervals.

Beyond financial data, LSTM finds applications in healthcare and activity classification. Uddin [39] showcased how LSTM models, when fed data from wearable healthcare sensors via edge devices, could adeptly classify twelve distinct human activities. The research underscored the robustness of the LSTM approach, outperforming traditional models. In [40], another application of LSTM introduced a deep learning model for Automatic Modulation Classification. This model integrates the capabilities of the residual neural network (ResNet) with LSTM, culminating in a receiver that can distinguish ten modulation techniques. From the aforementioned applications, it is evident that LSTM stands as a formidable tool for handling time series data.

2.2. LSTM Formulation and Structure

As illustrated in Figure 1, an LSTM cell takes an input series represented as $i = [i_1, i_2, \dots, i_T]$, where T denotes the length of the time series. This series is introduced into the network, resulting in the formation of a hidden sequence denoted as $h = [h_1, h_2, \dots, h_T]$. Each network activation unit manages the sequential data iteratively over time, converting the input series to the hidden series in accordance with specific mathematical equations as follows:

$$g_t = \sigma(W_{gi}i_t + W_{gh}h_{t-1} + b_g) \quad (1)$$

$$m_t = \sigma(W_{mi}i_t + W_{mh}h_{t-1} + b_m) \quad (2)$$

$$\tilde{j}_t = \tanh(W_{ji}i_t + W_{jh}h_{t-1} + b_j) \quad (3)$$

$$n_t = \sigma(W_{ni}i_t + W_{nh}h_{t-1} + b_n) \quad (4)$$

$$j_t = (g_t \odot j_{t-1}) + (m_t \odot \tilde{j}_t) \tag{5}$$

$$h_t = n_t \odot \tanh(j_t) \tag{6}$$

where h_t denotes the output of the hidden state at the current time step. The forget, input, and output gates are represented by symbols g , m , and n , respectively. The candidate cell vector is depicted as \tilde{j}_t , whereas j represents the cell activation vector. The weight matrices and bias vector are denoted by the terms W and b , respectively. The symbol \odot is used to indicate the Hadamard product. The activation functions for sigmoid and hyperbolic tangent are represented by the symbols σ and \tanh , respectively.

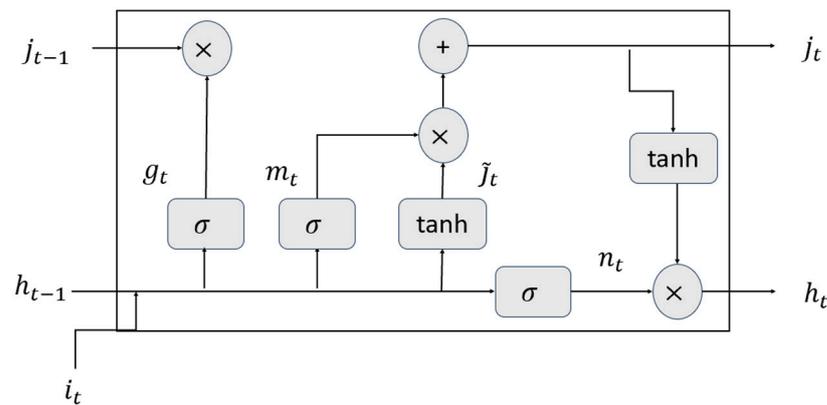


Figure 1. The structure of LSTM cell.

2.3. CNNs for Classification

CNNs are a type of deep learning model that are primarily used for analyzing visual imagery [41,42]. They are especially effective due to their ability to capture spatial dependencies in the data by applying filters [43]. Figure 2 shows the structure of a CNN for image classification.

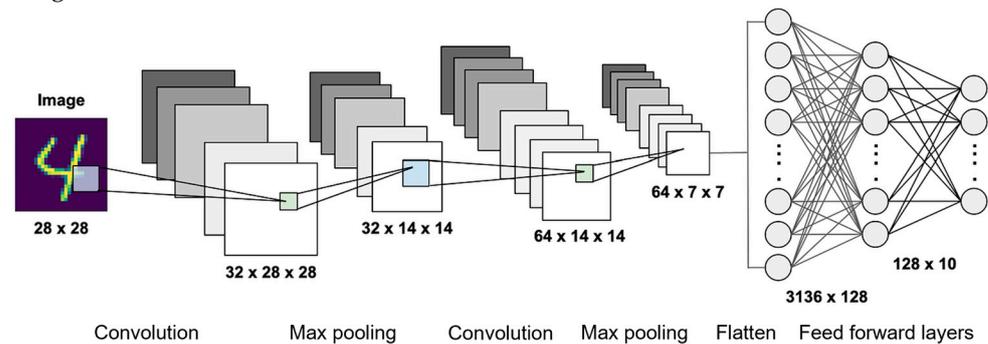


Figure 2. Structure of CNN for image classification.

While CNNs are widely acknowledged for their proficiency in image processing, their utility extends beyond this domain. They are also capable of handling time series classification, thus demonstrating their versatility in handling different types of data [44,45]. CNNs can handle sequential data by treating it as a one-dimensional “image” and applying convolutional filters to extract local and global patterns, which can often be beneficial in time series analysis.

The application of CNNs in time series data analysis offers several benefits. One of these is the automatic extraction and learning of features from raw data. This means that instead of manually engineering features from the data (a process that can be labor-intensive and prone to human bias), CNNs have the ability to automatically identify and learn relevant features through training. This automation makes CNNs an efficient tool

for handling complex data, as they can identify patterns and relationships that might be overlooked by human analysts. Another advantage of CNNs is their ability to reduce dimensionality. In complex datasets, there can be a large number of features (or dimensions), not all of which are necessarily useful or relevant. CNNs, through the use of convolutional and pooling layers, can distill these high-dimensional data into a lower-dimensional form while still preserving the most informative features. This process simplifies the data, making them easier to work with and reducing the computational resources required. Lastly, CNNs are well equipped to handle large datasets. They are designed to process input data in small, overlapping chunks (or windows), which makes them scalable and capable of handling large amounts of data without a proportional increase in computational requirements. This is particularly useful in scenarios where data are abundant and cannot be efficiently processed in one attempt. These attributes make CNNs a particularly good fit for tasks involving temporal dependencies.

3. Proposed Ensemble Technique

The ensemble neural network architecture, which harmoniously orchestrates multiple neural networks, plays a crucial role in bolstering the efficiency of classification systems [46]. Its versatile applications span across a myriad of classification tasks, demonstrating its profound significance. This particular framework integrates the predictions drawn from a plethora of individual models, often referred to as base learners or weak classifiers, each trained independently. The collective inference from these diverse models culminates in the final classification verdict, embodying the strength of collaborative decision-making.

Employing LSTM and CNN in the classification of ECG signals presents a plethora of benefits. LSTM, with its distinct capability to identify and encapsulate long-term dependencies in sequential data, is exceptionally suited for interpreting ECG signals, which are inherently time series data. In contrast, CNNs are highly effective at discerning spatial patterns and hierarchies, thereby enabling them to reveal subtle features within ECG data that could indicate specific heart conditions. The amalgamation of LSTM's expertise in temporal sequence recognition and CNN's skill in spatial pattern identification offers a robust and highly accurate approach for ECG signal analysis. This combined methodology greatly enhances the precision of cardiac anomaly diagnosis, the personalization of treatments, and overall improvement in patient care within the field of cardiology.

Inspired by these considerations, we propose our ensemble architecture, which strategically leverages the strengths of LSTM and CNN models simultaneously. This novel approach aims to harness the power of both these methodologies, optimizing their individual benefits in a cooperative manner. Our proposed approach involves one-hot encoding the outputs from three individual classifiers, integrating them with the raw time series signal as well as transformed signals, and subsequently using them in the final classifier—an LSTM–CNN. The ensemble classifier structure for ECG signal classification is graphically illustrated in Figure 3. The fundamental principle underlying our ensemble architecture is the belief that inherent diversity among individual models substantially enhances the system's accuracy, resilience, and generalization capabilities. The following discussion expounds upon the benefits of our ensemble methodology.

Our proposed ensemble architecture brings significant advantages to the table when applied to time series classification, especially concerning the analysis of ECG signals. ECG signals, being dynamic and intricate in nature, are a rich source of information about the electrical activity of the heart. Our ensemble structure acts as a powerful tool to tackle the inherent variability and noise that are frequently encountered in ECG data. By combining predictions from a multitude of models, this ensemble approach manages to harness a broader range of patterns and is better equipped to adapt to the diverse characteristics of ECG signals. This results in a notable enhancement in the accuracy of ECG signal classification, which can be invaluable for diagnosing heart-related disorders, refining prosthetic control, and improving the outcomes of rehabilitation programs. Moreover,

the robustness of our ensemble architecture against noisy ECG signals ensures a reliable analysis, considerably reducing the impact of signal artifacts on classification results.

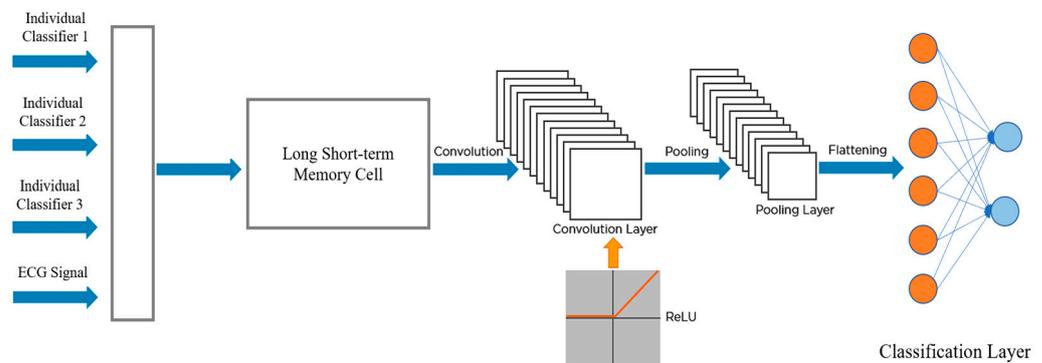


Figure 3. The structure of the proposed classifier for EMG signal classification.

It is important to acknowledge certain challenges inherent to the proposed ensemble neural network, which integrates LSTM and CNN architectures. First, while the design captures both spatial and temporal nuances of ECG signals, allowing it to manage data with minor to moderate noise levels, it can falter when confronted with highly noisy data. Such significant noise can mask pivotal features, potentially compromising the model's ability to discern patterns. To circumvent this limitation, implementing noise reduction techniques before feeding data into the model is recommended [47]. Second, consistent with the nature of ensemble methods, the computational demands of this model exceed those of individual neural networks. Third, due to its intricate architecture, the model's performance might not meet the expected benchmarks when faced with an exceedingly limited dataset. However, data augmentation presents itself as an effective countermeasure to this challenge.

4. Pre-Classification Signal Processing

The dataset employed for this research, consisting of ECG recordings obtained using the AliveCor device, was generously made available by AliveCor [48]. The training subset consists of 8528 single-lead ECG recordings, each lasting between 9 s and slightly over 60 s. In comparison, the test subset houses 3658 ECG recordings of similar lengths. These ECG recordings were sampled at a rate of 300 Hz and were subject to band-pass filtering using the AliveCor device. In Figure 4, we present visual samples drawn from our comprehensive dataset, capturing the stark differences between atrial fibrillation and normal cardiac signals. Atrial fibrillation, a widely recognized cardiac arrhythmia, exhibits specific characteristics on the ECG, notably its rapid and irregular beats. In juxtaposition, the 'normal' cardiac signal epitomizes a rhythmic and systematic heart activity. However, even this 'normal' signal occasionally presents irregular anomalies (shown in Figure 4), complicating the classification process. Such intricacies make classification challenging, if not unfeasible, with rudimentary models. Hence, there is a compelling necessity to employ more robust and powerful classifiers.

To enhance the performance of the classifier, we implement a feature extraction process. The choice of features to extract is guided by the computation of spectrograms, which we later employ as input for our deep learning network. Figure 5 displays the spectrograms of both categories of signals in our dataset.

We then transform the spectrograms into one-dimensional signals using Time–Frequency (TF) moments. Each TF moment gleams specific information from the spectrogram, thus serving as a distinctive one-dimensional feature that can be inputted into the LSTM network. For this investigation, we focus on its spectral entropy. Spectral entropy is a measure of the flatness or “spikiness” of a signal's spectrum [49,50]. A signal with a “spiky”

spectrum (analogous to a sum of sinusoids) exhibits low spectral entropy, whereas a signal with a flat spectrum (such as white noise) shows high spectral entropy.

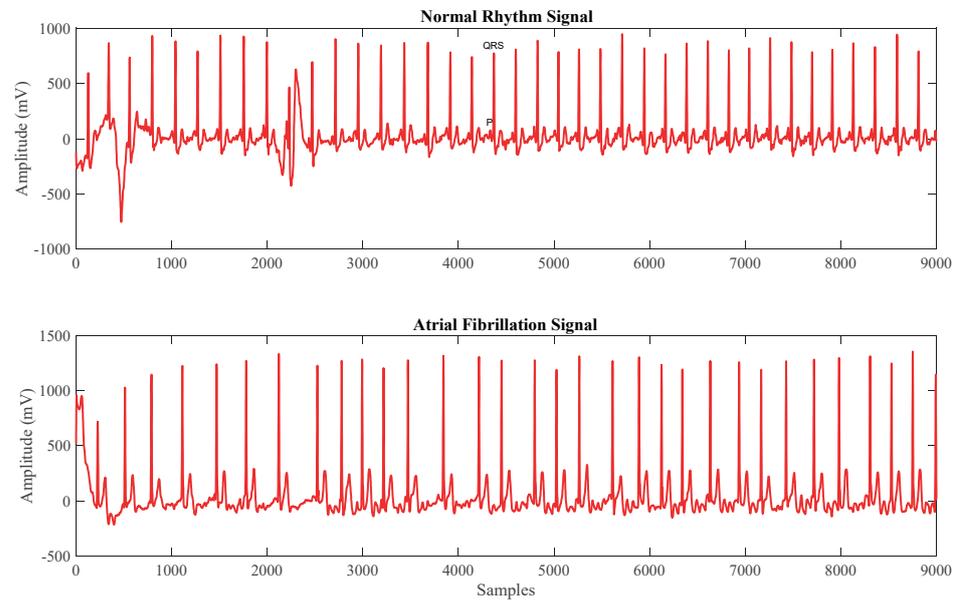


Figure 4. Raw samples from the ECG dataset.

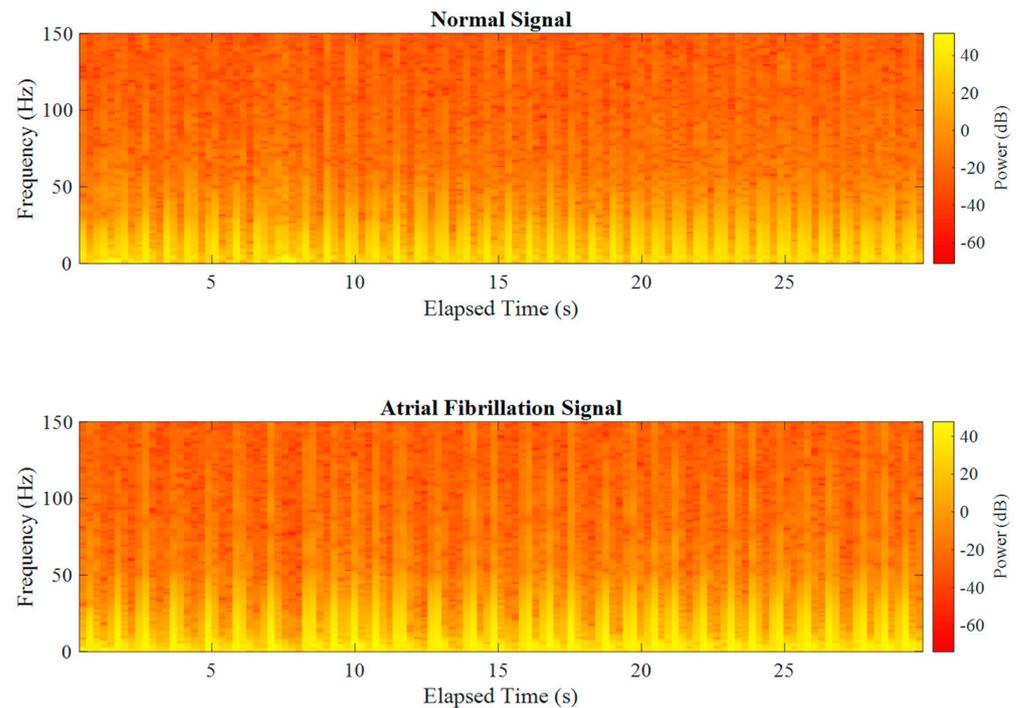


Figure 5. The calculated spectrograms of two signals in the ECG dataset.

We derive the spectral entropy based on a power spectrogram, utilizing 255 time windows for the computation, similar to the approach adopted for instantaneous frequency estimation. Figure 6 portrays the spectral entropy for each category of signal in our dataset. These refined features and the adopted signal processing methodology significantly contribute to the enhanced classification efficacy, as shown in the next section.

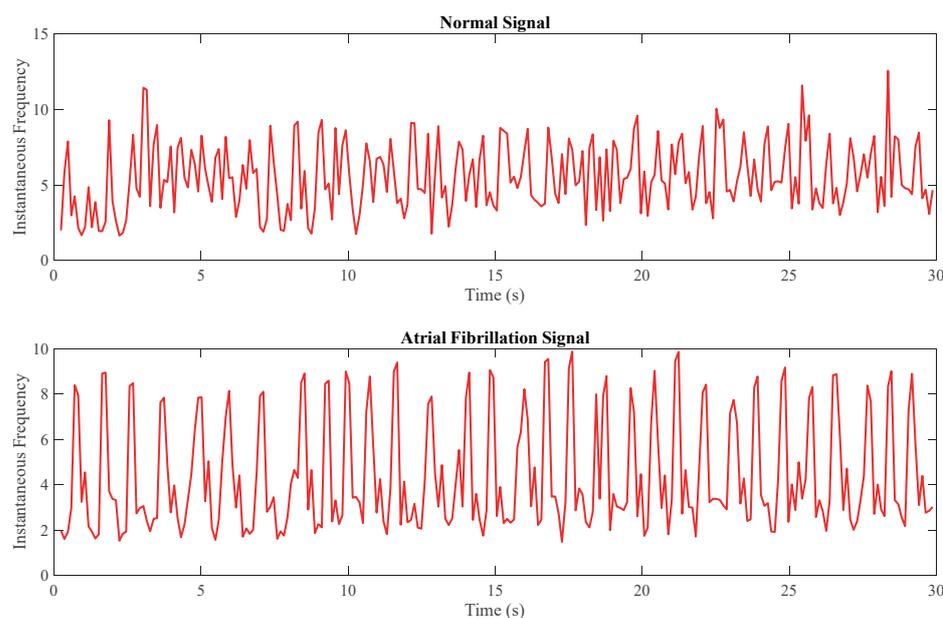


Figure 6. The calculated spectral entropy of two signals in the ECG dataset.

5. Numerical Results and Comparison

In this section, we provide a thorough evaluation of our proposed model's performance, laying out an in-depth comparison against two simulations. The first simulation employed for comparison is a conventional LSTM model, which uses raw, unprocessed ECG signals. The primary aim of selecting this model is to assess the inherent capacity of the LSTM model to discern and learn from patterns within raw data and compare this ability with that of our proposed model. The second comparison method involves another conventional LSTM model. In contrast to the first method, this model utilizes ECG signals that have already undergone a preprocessing stage along with raw data. The specific preprocessing techniques applied were detailed and explained in Section 4.

5.1. LSTM for Raw Time Series

The LSTM architecture implemented in this scenario starts with a sequence input layer, specifically engineered to handle an input sequence array of one dimension. This input layer is subsequently connected to an LSTM layer comprising 100 hidden units. A dropout layer, with a dropout rate of 0.2, is integrated next to mitigate overfitting. Subsequently, the network includes two fully connected layers, with 20 and 2 neurons, respectively, that classify the learned features. The outputs are then passed through a SoftMax layer that normalizes them into probabilities, before finally reaching a classification layer.

Figure 7 provides a visual representation of the model's accuracy and corresponding loss for this particular case. This figure elucidates the trajectory of the accuracy and its associated loss function throughout the training process. It reveals an apparent issue: the model is struggling to extract significant features from the raw signal. This highlights an essential requirement for preprocessing the signal prior to inputting the data into the LSTM network.

Figure 8 presents the classification results obtained from testing data. In the figure presented, the vertical axis, labeled "True Classes", delineates the actual categories or ground truth of the data samples. In contrast, the horizontal axis, termed "Predicted Class", captures the classifications as perceived by the model. Together, these axes offer a visual comparison between the model's predictions and the real labels, enabling an immediate assessment of classification accuracy and areas of potential discrepancy. It is evident that, despite numerous iterations, the LSTM model struggles to classify the raw signals effectively, revealing a lack of comprehension of the underlying function inherent in the data. This is further demonstrated by the fact that both true positives and true negatives

account for less than 50% of the results. These observations highlight a pressing need for signal processing before feeding data to the LSTM model.

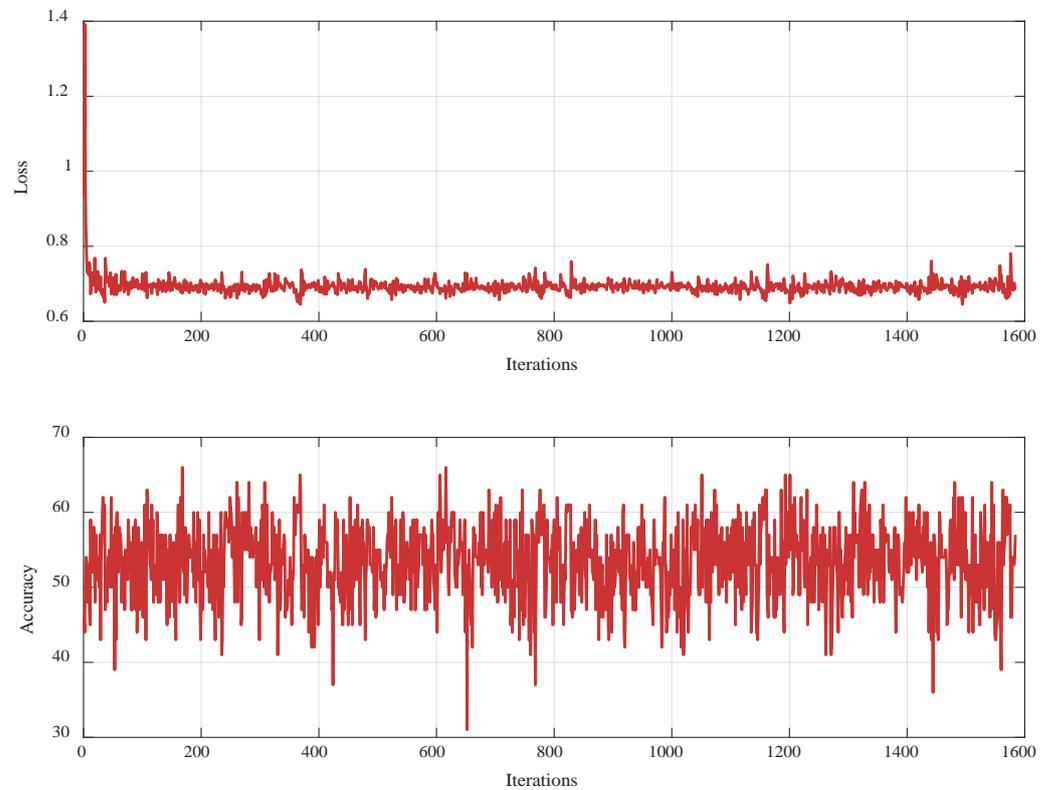


Figure 7. The accuracy and loss of LSTM network applied to the raw ECG dataset.

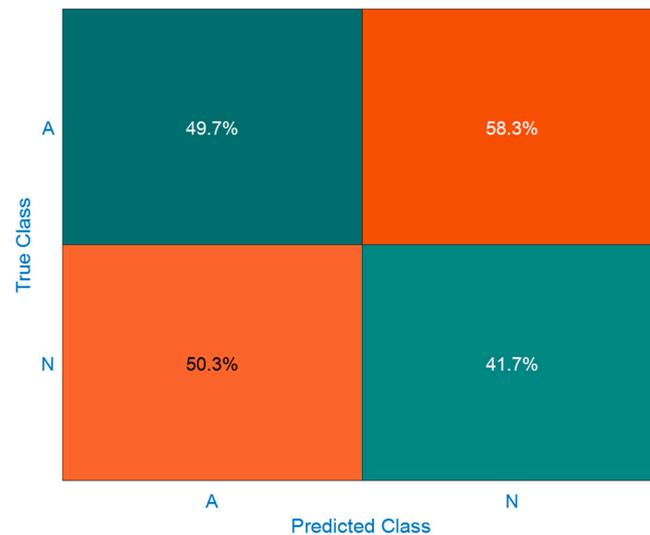


Figure 8. The classification results of LSTM applied to the raw ECG dataset where the vertical axis represents 'True Classes' (actual dataset labels) and the horizontal axis displays the 'Predicted Class' as determined by the model.

5.2. LSTM to Preprocess Time Series

Here, maintaining the previously defined neural network architecture, the preprocessed ECG data (delineated in Section 4) is concatenated with the raw ECG signals to form an augmented dataset. This composite data serves as the input for the LSTM, allowing the model to concurrently leverage the enhanced features of the preprocessed data and the untouched, subtle details present in the raw data.

Figure 9 illustrates the evolution of the model’s accuracy and loss function throughout the training phase for this case, which incorporates both raw and preprocessed data. As can be observed, significant shifts in both metrics occur, suggesting the successful adaptation of the model in understanding the underlying function of the data. This is in stark contrast to the previous scenario as evidenced in Figure 7, where the LSTM model struggled with raw data alone. Upon concluding the training phase, the classification efficacy of this case is evaluated on the test data. Figure 10 offers a graphical depiction of these results. Upon comparing these outcomes with those illustrated in Figure 8, there is a substantial improvement observed.

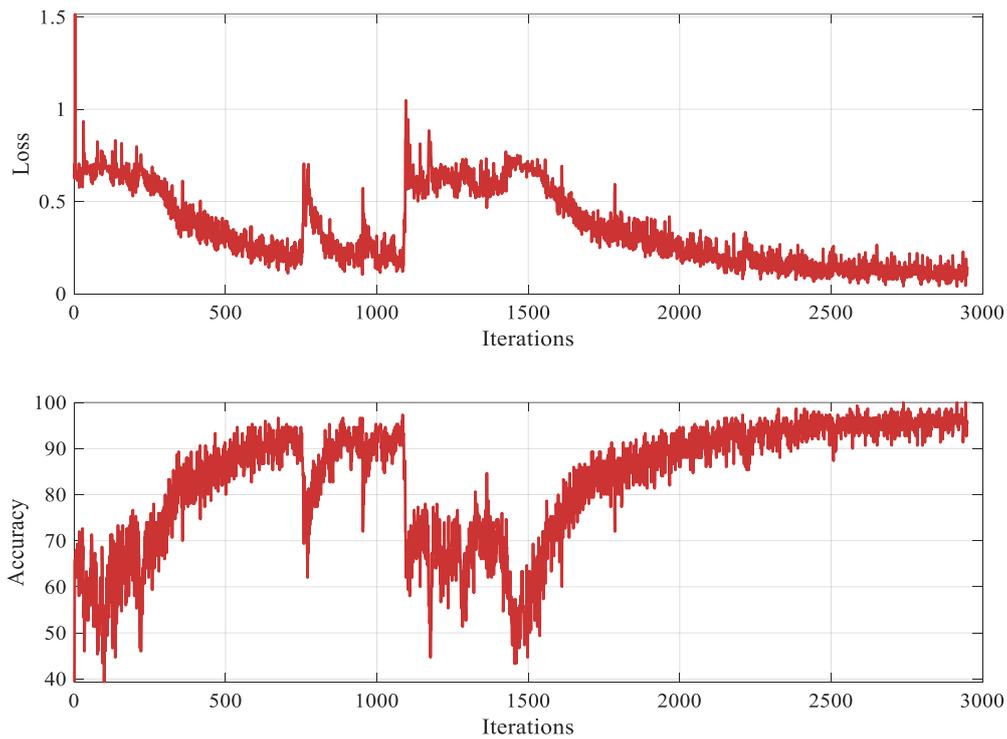


Figure 9. The accuracy and loss of LSTM network applied to the processed ECG dataset, showcasing the evolution of accuracy and loss over training iterations.

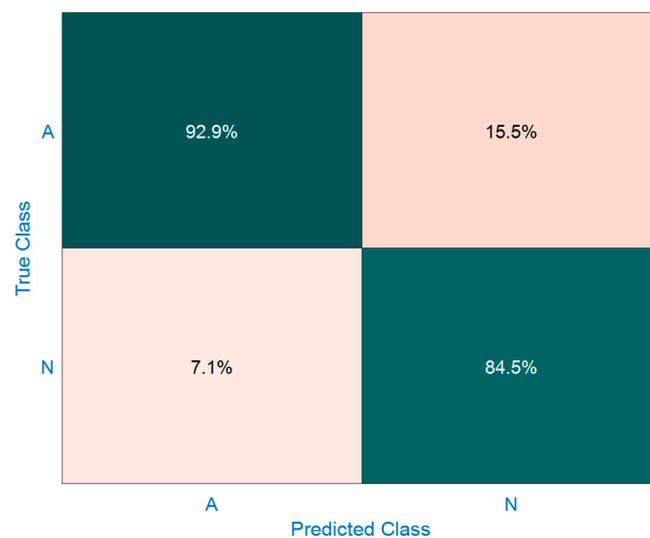


Figure 10. The classification results of LSTM applied to the processed ECG dataset.

5.3. Proposed Approach

In our proposed methodology, we implement three distinct neural networks, each characterized by its own unique design. Following this, the outputs of these networks are amalgamated with both raw and preprocessed data, which then act as the input for a culminating network. The first classifier in our framework utilizes a bidirectional LSTM layer housing 100 hidden units. Subsequently, this forwards only the final sequence output to a two-neuron fully connected layer. Its terminal layers are composed of a regularization layer and a classification layer. The second classifier is designed with an LSTM layer containing 80 hidden units, which feeds into a fully connected layer with two neurons. The third classifier integrates an LSTM layer with 80 hidden units and a connected layer with a pair of neurons. This classifier concludes with a regularization layer followed by its classification layer. However, it is important to note that the number of layers and neurons in the proposed structure can be adjusted based on the specific application at hand.

Our concluding network initiates with a sequence input layer tailored for the input dataset, succeeded by a 1D convolution layer equipped with 96 filters, a max pooling layer with a pool size of 3, and a LSTM layer having 25 hidden units. The architecture is rounded off with a two-neuron fully connected layer and a final classification layer. Regarding training specifics, every network employs the Adam optimizer, with a cap of 30 epochs, a mini-batch size of 150, and an inaugural learning rate of 0.01. To counteract oversized gradients, a gradient threshold of 1 is set in place.

As displayed in Figure 11, during the training phase, the model achieved convergence within a limited number of steps, marking a significant improvement in learning efficiency compared to the conventional LSTM model (see Figures 7 and 9). The classification results of the proposed LSTM–CNN ensemble model are presented in Figure 12. The ensemble model exceeded the performance of the conventional LSTM model by a substantial margin, confirming its superior capabilities in ECG signal classification.

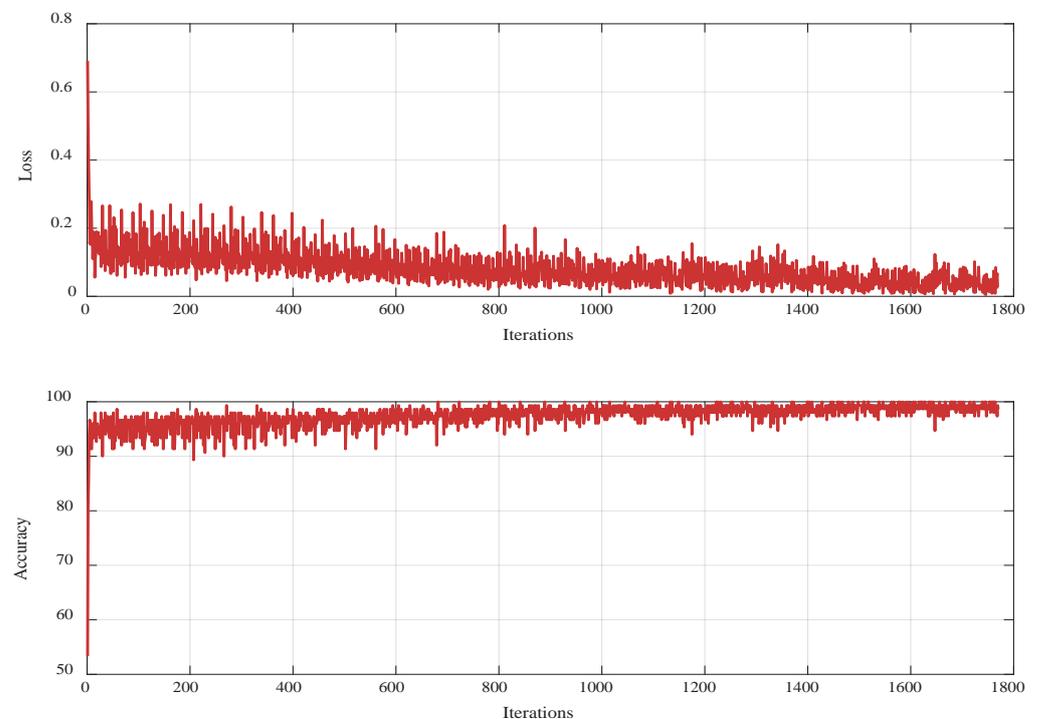


Figure 11. The accuracy and loss of the proposed method applied to the processed ECG dataset, highlighting the swift convergence of accuracy and loss over training iterations.

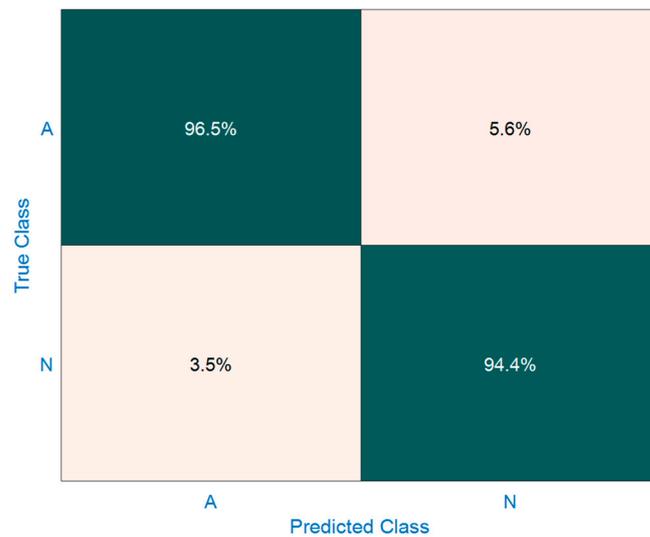


Figure 12. The classification results of the proposed method applied to the processed ECG dataset.

When juxtaposing Figures 8, 10 and 12, it becomes evident that our proposed approach, combined with preprocessing, induces a substantial improvement in the classification outcomes. The stark contrast in performance is especially noticeable when comparing the average accuracy rates. While Figures 8 and 10 demonstrate the limitations of traditional and raw data methods, Figure 12 showcases the efficacy of our method. In this case, the proposed method’s average accuracy is markedly high, achieving a score of 95.45%. These results underscore the value of implementing a combined strategy of data preprocessing and advanced network architecture in significantly enhancing the classification accuracy of ECG signals.

5.4. Comparison with the Bidirectional GRU Network Model

In this section, we contrast our approach with a recent method introduced in [51]. This method employs a deep RNN, particularly leveraging the Gated Recurrent Unit (GRU) in a bidirectional configuration. The aforementioned study reports that their bidirectional GRU model, a fusion of RNN and GRU in a bidirectional setting, delivers impressive classification accuracy. To ensure a robust comparison, we utilized their model on our dataset, ensuring both techniques are evaluated using identically preprocessed data. The subsequent figure illustrates the achieved accuracy.

A comparative analysis between these results (Figure 13) and ours, as depicted in Figure 12, distinctly demonstrates the superior performance of our method. This enhancement is attributed to the advantageous pattern recognition capabilities of CNNs in our structure. Furthermore, our approach, due to its ensemble structure, exhibits resilience against overfitting—a pervasive issue in machine learning. Conversely, the bidirectional GRU model is predisposed to overfitting and necessitates meticulous tuning for real-world applications.

For a thorough understanding of the comparative performance of all the methods under study, we delineate their sensitivity, specificity, and accuracy, of which their formulations are given by:

$$\begin{aligned}
 \text{Sensitivity} &= \frac{TP}{TP+FN} \\
 \text{Specificity} &= \frac{TN}{TN+FP} \\
 \text{Accuracy} &= \frac{TP+TN}{TP+FN+TN+FP}
 \end{aligned}
 \tag{7}$$

We refer to the cases with atrial fibrillation as ‘positive’, denoted by (P), and to those without this condition (normal cases) as ‘negative’, represented by (N). Consequently, TP denotes the number of True Positives, TN represents the number of True Negatives. Additionally, FN signifies the number of False Negatives, and FP stands for the number of False Positives. These metrics provide different insights into the performance of a binary

classifier. For instance, Sensitivity shows how well the model identifies positive cases, while Specificity indicates how well the model identifies negative cases. Accuracy provides an overall measure of how often the model is correct, regardless of the class.

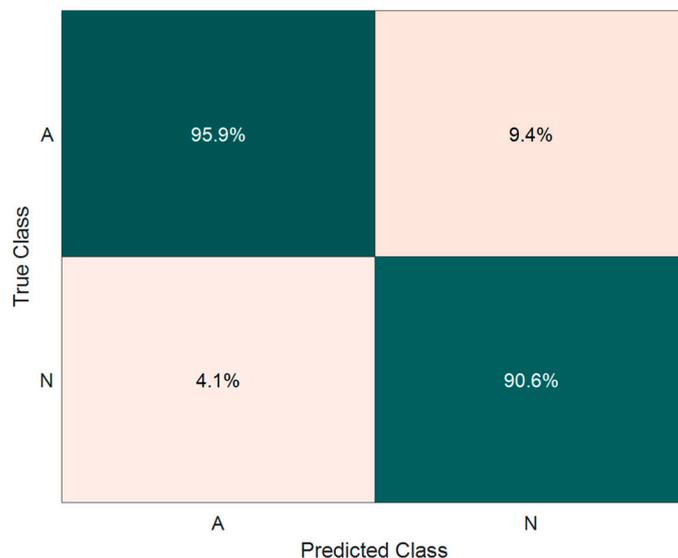


Figure 13. The classification results of the bidirectional GRU network applied to the processed ECG dataset.

Table 1 summarizes the results across all methods applied in this study. As detailed in the table, our proposed approach consistently surpasses other methods in performance metrics. Notably, even when juxtaposed with the bidirectional GRU—a sophisticated and advanced model specifically tailored for such datasets—our method is superior.

Table 1. Numerical comparison of applied methods.

Method	Sensitivity	Specificity	Accuracy
LSTM applied to raw data	46.02	45.33	45.70
LSTM applied to preprocessed data	85.70	92.25	88.70
The proposed method applied to preprocessed data	94.52	96.42	95.45
Bidirectional GRU applied to preprocessed data	91.07	95.67	93.25

Looking ahead, there is a pressing need for future research to focus on advanced noise reduction techniques, specifically tailored for ECG data. Such techniques are vital in filtering out extraneous disturbances, which frequently obscure ECG readings and pose challenges in signal interpretation. Beyond noise reduction, the integration of data augmentation methods stands as a promising avenue. By artificially expanding and diversifying the dataset, these methods enable models to train on a broader spectrum of cardiac patterns. This not only bolsters the model’s predictive accuracy, but also enhances its generalization capabilities across varied and unseen ECG patterns. Moreover, considering the interdisciplinary nature of the problem, collaborations between biomedical engineers, data scientists, and cardiologists could yield novel insights and methodologies. By amalgamating these strategies and fostering collaborative efforts, the path is paved toward significantly elevating the reliability and precision of ECG signal classification. Such advancements will undoubtedly result in systems that are both robust and consistently accurate across diverse clinical and real-world settings.

6. Conclusions

An advanced ensemble methodology has been introduced, specifically crafted for ECG signal classification. By integrating the strengths of LSTM and CNN models, a significant enhancement in classification accuracy has been achieved. The architecture and methodology have been scrupulously detailed. Additionally, the preprocessing steps implemented to transform raw ECG signals into a classification-ready format have been outlined. Frequency analysis was performed on the raw time series, and spectral entropy was calculated during the preprocessing stage, serving as an integral part of the classification process. In the results section, we made a comparison between the LSTM model trained on raw data and the one trained on preprocessed data, as well as our proposed ensemble architecture. This comparison underscores the vital role of data preprocessing and the efficacy of the proposed ensemble approach in enhancing model performance. The statistical outcomes have underscored the superiority of our approach. Our proposed technique has consistently achieved an accuracy exceeding 94%. By integrating this ensemble into wearable devices, it can offer real-time cardiac monitoring and instantaneous alerts for abnormal rhythms. As telehealth gains momentum, such a system could bolster remote patient management and diagnostic accuracy. Strategic collaborations with medical device manufacturers and health-tech startups can further drive innovation, while pilot testing in select healthcare facilities ensures the model's practical viability. Although our ensemble approach has yielded considerable advancements over the standard LSTM classifier, we acknowledge the perpetual potential for further research and optimization. Given the inherent dynamism and complexity of ECG data, future research should delve into advanced noise reduction techniques to filter out extraneous disturbances that often cloud ECG readings. Additionally, adopting data augmentation methods can enrich the dataset, allowing models to train on varied cardiac patterns, thereby enhancing their generalization capabilities. By amalgamating these strategies, we can significantly elevate the reliability of ECG signal classification, making it more robust and accurate in diverse settings.

Author Contributions: Conceptualization, N.S.A., H.J., Q.Y., S.B. and I.M.; methodology, N.S.A., H.J., Q.Y., S.B. and I.M.; software, N.S.A., H.J., Q.Y., S.B. and I.M.; validation, N.S.A., H.J., Q.Y., S.B. and I.M.; writing—original draft preparation, N.S.A., H.J., Q.Y., S.B. and I.M.; writing—review and editing, N.S.A., H.J., Q.Y., S.B. and I.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research work was funded by Institutional Fund Projects under grant no. (IFPIP: 244-247-1443). The authors gratefully acknowledge technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia.

Data Availability Statement: Not Applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Del Re, D.P.; Amgalan, D.; Linkermann, A.; Liu, Q.; Kitsis, R.N. Fundamental Mechanisms of Regulated Cell Death and Implications for Heart Disease. *Physiol. Rev.* **2019**, *99*, 1765–1817. [[CrossRef](#)] [[PubMed](#)]
2. Coffey, S.; Roberts-Thomson, R.; Brown, A.; Carapetis, J.; Chen, M.; Enriquez-Sarano, M.; Zühlke, L.; Prendergast, B.D. Global Epidemiology of Valvular Heart Disease. *Nat. Rev. Cardiol.* **2021**, *18*, 853–864. [[CrossRef](#)] [[PubMed](#)]
3. Yan, Y.; Zhang, J.-W.; Zang, G.-Y.; Pu, J. The Primary Use of Artificial Intelligence in Cardiovascular Diseases: What Kind of Potential Role Does Artificial Intelligence Play in Future Medicine? *J. Geriatr. Cardiol. JGC* **2019**, *16*, 585.
4. Khan, M.A.; Algarni, F. A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS. *IEEE Access* **2020**, *8*, 122259–122269. [[CrossRef](#)]
5. Butt, F.S.; La Blunda, L.; Wagner, M.F.; Schäfer, J.; Medina-Bulo, I.; Gómez-Ullate, D. Fall Detection from Electrocardiogram (Ecg) Signals and Classification by Deep Transfer Learning. *Information* **2021**, *12*, 63. [[CrossRef](#)]
6. Murat, F.; Yildirim, O.; Taló, M.; Baloglu, U.B.; Demir, Y.; Acharya, U.R. Application of Deep Learning Techniques for Heartbeats Detection Using ECG Signals-Analysis and Review. *Comput. Biol. Med.* **2020**, *120*, 103726. [[CrossRef](#)] [[PubMed](#)]
7. Ahmed, Z.; Mohamed, K.; Zeeshan, S.; Dong, X. Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine. *Database* **2020**, *2020*, baaa010. [[CrossRef](#)]

8. Galbusera, F.; Casaroli, G.; Bassani, T. Artificial Intelligence and Machine Learning in Spine Research. *JOR Spine* **2019**, *2*, e1044. [[CrossRef](#)]
9. Khalid, H.; Hussain, M.; Al Ghamdi, M.A.; Khalid, T.; Khalid, K.; Khan, M.A.; Fatima, K.; Masood, K.; Almotiri, S.H.; Farooq, M.S. A Comparative Systematic Literature Review on Knee Bone Reports from Mri, X-rays and Ct Scans Using Deep Learning and Machine Learning Methodologies. *Diagnostics* **2020**, *10*, 518. [[CrossRef](#)]
10. Seifert, R.; Weber, M.; Kocakavuk, E.; Rischpler, C.; Kersting, D. *Artificial Intelligence and Machine Learning in Nuclear Medicine: Future Perspectives*; Elsevier: Amsterdam, The Netherlands, 2021; Volume 51, pp. 170–177.
11. Nagendra, H.; Mukherjee, S.; Kumar, V. Application of Wavelet Techniques in ECG Signal Processing: An Overview. *Int. J. Eng. Sci. Technol. IJEST* **2011**, *3*, 7432–7443.
12. He, H.; Tan, Y. Automatic Pattern Recognition of ECG Signals Using Entropy-Based Adaptive Dimensionality Reduction and Clustering. *Appl. Soft Comput.* **2017**, *55*, 238–252. [[CrossRef](#)]
13. Libbrecht, M.W.; Noble, W.S. Machine Learning Applications in Genetics and Genomics. *Nat. Rev. Genet.* **2015**, *16*, 321–332. [[CrossRef](#)] [[PubMed](#)]
14. Whalen, S.; Schreiber, J.; Noble, W.S.; Pollard, K.S. Navigating the Pitfalls of Applying Machine Learning in Genomics. *Nat. Rev. Genet.* **2022**, *23*, 169–181. [[CrossRef](#)] [[PubMed](#)]
15. Yousefpour, A.; Jahanshahi, H.; Bekiros, S. Optimal Policies for Control of the Novel Coronavirus Disease (COVID-19) Outbreak. *Chaos Solitons Fractals* **2020**, *136*, 109883. [[CrossRef](#)]
16. Houssein, E.H.; Kilany, M.; Hassanien, A.E. ECG Signals Classification: A Review. *Int. J. Intell. Eng. Inform.* **2017**, *5*, 376–396. [[CrossRef](#)]
17. Mincholé, A.; Camps, J.; Lyon, A.; Rodríguez, B. Machine Learning in the Electrocardiogram. *J. Electrocardiol.* **2019**, *57*, S61–S64. [[CrossRef](#)] [[PubMed](#)]
18. Xiong, P.; Lee, S.M.-Y.; Chan, G. Deep Learning for Detecting and Locating Myocardial Infarction by Electrocardiogram: A Literature Review. *Front. Cardiovasc. Med.* **2022**, *9*, 860032. [[CrossRef](#)]
19. Li, F.; Chang, H.; Jiang, M.; Su, Y. A Contrastive Learning Framework for ECG Anomaly Detection. In Proceedings of the 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 15–17 April 2022; pp. 673–677.
20. Hagiwara, Y.; Fujita, H.; Oh, S.L.; Tan, J.H.; San Tan, R.; Ciaccio, E.J.; Acharya, U.R. Computer-Aided Diagnosis of Atrial Fibrillation Based on ECG Signals: A Review. *Inf. Sci.* **2018**, *467*, 99–114. [[CrossRef](#)]
21. Yu, Y.; Si, X.; Hu, C.; Zhang, J. A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures. *Neural Comput.* **2019**, *31*, 1235–1270. [[CrossRef](#)]
22. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaria, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)]
23. Cheng, M.; Sori, W.J.; Jiang, F.; Khan, A.; Liu, S. Recurrent Neural Network Based Classification of ECG Signal Features for Obstruction of Sleep Apnea Detection. In Proceedings of the 2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC), Guangzhou, China, 21–24 July 2017; Volume 2, pp. 199–202.
24. Naul, B.; Bloom, J.S.; Pérez, F.; van der Walt, S. A Recurrent Neural Network for Classification of Unevenly Sampled Variable Stars. *Nat. Astron.* **2018**, *2*, 151–155. [[CrossRef](#)]
25. Welch, R.L.; Ruffing, S.M.; Venayagamoorthy, G.K. Comparison of Feedforward and Feedback Neural Network Architectures for Short Term Wind Speed Prediction. In Proceedings of the 2009 International Joint Conference on Neural Networks, Atlanta, GA, USA, 14–19 June 2009; pp. 3335–3340.
26. Sundermeyer, M.; Oparin, I.; Gauvain, J.-L.; Freiberger, B.; Schlüter, R.; Ney, H. Comparison of Feedforward and Recurrent Neural Network Language Models. In Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 8430–8434.
27. Kwon, D.-H.; Kim, J.-B.; Heo, J.-S.; Kim, C.-M.; Han, Y.-H. Time Series Classification of Cryptocurrency Price Trend Based on a Recurrent LSTM Neural Network. *J. Inf. Process. Syst.* **2019**, *15*, 694–706.
28. Tan, H.X.; Aung, N.N.; Tian, J.; Chua, M.C.H.; Yang, Y.O. Time Series Classification Using a Modified LSTM Approach from Accelerometer-Based Data: A Comparative Study for Gait Cycle Detection. *Gait Posture* **2019**, *74*, 128–134. [[CrossRef](#)] [[PubMed](#)]
29. Hochreiter, S.; Schmidhuber, J. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [[CrossRef](#)]
30. Men, L.; Ilk, N.; Tang, X.; Liu, Y. Multi-Disease Prediction Using LSTM Recurrent Neural Networks. *Expert Syst. Appl.* **2021**, *177*, 114905. [[CrossRef](#)]
31. Yang, B.; Yin, K.; Lacasse, S.; Liu, Z. Time Series Analysis and Long Short-Term Memory Neural Network to Predict Landslide Displacement. *Landslides* **2019**, *16*, 677–694. [[CrossRef](#)]
32. Weerakody, P.B.; Wong, K.W.; Wang, G.; Ela, W. A Review of Irregular Time Series Data Handling with Gated Recurrent Neural Networks. *Neurocomputing* **2021**, *441*, 161–178. [[CrossRef](#)]
33. Hashemi, R.; Brigode, P.; Garambois, P.-A.; Javelle, P. How Can We Benefit from Regime Information to Make More Effective Use of Long Short-Term Memory (LSTM) Runoff Models? *Hydrol. Earth Syst. Sci.* **2022**, *26*, 5793–5816. [[CrossRef](#)]
34. Fischer, T.; Krauss, C. Deep Learning with Long Short-Term Memory Networks for Financial Market Predictions. *Eur. J. Oper. Res.* **2018**, *270*, 654–669. [[CrossRef](#)]

35. Yan, H.; Ouyang, H. Financial Time Series Prediction Based on Deep Learning. *Wirel. Pers. Commun.* **2018**, *102*, 683–700. [[CrossRef](#)]
36. Sagheer, A.; Kotb, M. Time Series Forecasting of Petroleum Production Using Deep LSTM Recurrent Networks. *Neurocomputing* **2019**, *323*, 203–213. [[CrossRef](#)]
37. Cen, Z.; Wang, J. Crude Oil Price Prediction Model with Long Short Term Memory Deep Learning Based on Prior Knowledge Data Transfer. *Energy* **2019**, *169*, 160–171. [[CrossRef](#)]
38. Liu, Y. Novel Volatility Forecasting Using Deep Learning—Long Short Term Memory Recurrent Neural Networks. *Expert Syst. Appl.* **2019**, *132*, 99–109. [[CrossRef](#)]
39. Uddin, M.Z. A Wearable Sensor-Based Activity Prediction System to Facilitate Edge Computing in Smart Healthcare System. *J. Parallel Distrib. Comput.* **2019**, *123*, 46–53. [[CrossRef](#)]
40. Elsagheer, M.M.; Ramzy, S.M. A Hybrid Model for Automatic Modulation Classification Based on Residual Neural Networks and Long Short Term Memory. *Alex. Eng. J.* **2023**, *67*, 117–128. [[CrossRef](#)]
41. Fard, A.S.; Reutens, D.C.; Vegh, V. From CNNs to GANs for Cross-Modality Medical Image Estimation. *Comput. Biol. Med.* **2022**, *146*, 105556. [[CrossRef](#)]
42. Song, J.; Gao, S.; Zhu, Y.; Ma, C. A Survey of Remote Sensing Image Classification Based on CNNs. *Big Earth Data* **2019**, *3*, 232–254. [[CrossRef](#)]
43. Bodapati, S.; Bandarupally, H.; Shaw, R.N.; Ghosh, A. Comparison and Analysis of RNN-LSTMs and CNNs for Social Reviews Classification. In *Advances in Applications of Data-Driven Computing*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 49–59.
44. Sayadi, H.; Gao, Y.; Mohammadi Makrani, H.; Lin, J.; Costa, P.C.; Rafatirad, S.; Homayoun, H. Towards Accurate Run-Time Hardware-Assisted Stealthy Malware Detection: A Lightweight, yet Effective Time Series CNN-Based Approach. *Cryptography* **2021**, *5*, 28. [[CrossRef](#)]
45. Sadouk, L. CNN Approaches for Time Series Classification. In *Time Series Analysis-Data, Methods, and Applications*; IEEE: Rennes, France, 2019; Volume 5.
46. Gomes, H.M.; Barddal, J.P.; Enembreck, F.; Bifet, A. A Survey on Ensemble Learning for Data Stream Classification. *ACM Comput. Surv. CSUR* **2017**, *50*, 1–36. [[CrossRef](#)]
47. Chatterjee, S.; Thakur, R.S.; Yadav, R.N.; Gupta, L.; Raghuvanshi, D.K. Review of Noise Removal Techniques in ECG Signals. *IET Signal Process.* **2020**, *14*, 569–590. [[CrossRef](#)]
48. Clifford, G.D.; Liu, C.; Moody, B.; Li-Wei, H.L.; Silva, I.; Li, Q.; Johnson, A.; Mark, R.G. AF Classification from a Short Single Lead ECG Recording: The PhysioNet/Computing in Cardiology Challenge 2017. In Proceedings of the 2017 Computing in Cardiology (CinC), Rennes, France, 24–27 September 2017; pp. 1–4.
49. Han, N.C.; Muniandy, S.V.; Dayou, J. Acoustic Classification of Australian Anurans Based on Hybrid Spectral-Entropy Approach. *Appl. Acoust.* **2011**, *72*, 639–645. [[CrossRef](#)]
50. Liu, L.; Liu, B.; Huang, H.; Bovik, A.C. No-Reference Image Quality Assessment Based on Spatial and Spectral Entropies. *Signal Process. Image Commun.* **2014**, *29*, 856–863. [[CrossRef](#)]
51. Lynn, H.M.; Pan, S.B.; Kim, P. A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent Neural Networks. *IEEE Access* **2019**, *7*, 145395–145405. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.