ISNNI: 3007-3391

HEC Category: Y

Journal of Computers and Intelligent Systems

Vol. 4, No. 4, 2025, pages 244 - 255

Journal Homepage:

https://journals.iub.edu.pk/index.php/JCIS/

Enhanced Wearable ECG Monitoring with LSTM Autoencoder-Based Anomaly Detection

Ali Raza Abbasi[1], Samra Aslam[2], Saba Akram[3], Muhammad Akbar Urf Mujahid[4], Muhammad Asad[5]

[1]Departemnt of Computer Science, The Islamia University of Bahawalpur -63100, PUNJAB, Pakistan
aliabbasibwp12@gmail.com

[2]Departemnt of Computer Science,The Islamia University of Bahawalpur -63100, PUNJAB, Pakistan
samra.aslamofficial@gmail.com

[3]Departemnt of Computer Science,The Islamia University of Bahawalpur -63100, PUNJAB, Pakistan
akramsaba680@gmail.com

[4]Departemnt of Computer Science,The Islamia University of Bahawalpur -63100, PUNJAB, Pakistan
malikmujahid.ch82@gmail.com

[5]Departemnt of Artificial Intelligence,The Islamia University of Bahawalpur -63100, PUNJAB, Pakistan
asadblouch1994@gmail.com

ARTICLE INFO

Article History:

Received:	September	28, 2025
Revised:	September	30, 2025
Accepted:	October	26, 2025
Available Online:	October	30, 2025

Keywords:

Wearable Sensors
Anomaly Detection
LSTM Autoencoder
ECG
Machine Learning
Deep Learning

Classification Codes:

Funding:

This research received no specific grant from any funding agency in the public or not-for-profit sector.

ABSTRACT

Cardiovascular diseases (CVDs) remain as the leading cause of death worldwide as they cause close to 18 million deaths in a year. Early and precise diagnosis of heart defects like arrhythmia, ischemia, and heart failure is still an important consideration in minimizing the mortality rate and enhancing the quality of life. Smartwatches, chest straps and patches are wearable body sensors that turned out to be revolutionary in terms of continuous and non-invasive cardiac health monitoring. Such devices produce large amounts of Electrocardiogram (ECG) in real-time, which serves as a useful source of anomaly detection. Nonetheless, the proper analysis of these data is complicated by such issues as temporal complexity, variability due to the human activity, inconsistency of the sensor locations, and imbalance in the number of classes where the abnormal cardiac events are underrepresented significantly. The conventional anomaly detection methods (rule-based thresholds, statistical models, etc.) find it difficult to cope with those issues, which results in high false alarm rates and low clinical usability. In order to overcome these shortcomings, this paper proposes an anomaly detection model using a Long Short-Term Memory (LSTM) Autoencoder on ECG signal. The framework takes advantage of the temporal learning of LSTM networks in an encoder-decoder framework to learn normal cardiac patterns and recognize the anomalies based on the reconstruction error. PhysioNet ECG dataset was used as an evaluation dataset and preprocessing activities were applied to include data cleansing, data normalization, time-series segmentation, as well as imbalanced classes. To measure the performance of the offered approach, the following classical machine learning algorithms were implemented and tested on the basis of the following measurements: precision, recall, F1-score, ROC-AUC, and the confusion matrices: Local Outlier Factor (LOF), Elliptic Envelope, K-Nearest Neighbors (KNN), Support Vector Machine (SVM), K-Means Clustering and Logistic Regression.

As it can be seen with the use of experimental findings, the proposed LSTM AutoEncoder has a major advantage, as it works much better than traditional methods and provides 99.45% accuracy, with even better precision and recall. In contrast to supervised methods, the model can be trained solely on normal ECG data so it can be more easily scaled and can be applied in a real-world environment where there are limited labeled anomaly data. This research identifies the promise of unsupervised models based on the deep learning approach in improving the accuracy, reliability, and real-time capacity of the wearable cardiac monitoring system. Moreover, the results open the door to the incorporation of the custom deviation identification structures into telemedical systems so that interventions could be implemented in time and the burden on healthcare services could be lessened.

1. Introduction

Cardiovascular diseases (CVDs) are considered to be a significant health issue in the world, killing millions of people each year. Recent research has also reported that untimely diagnosis and paucity in terms of monitoring options remain the factors in the avoidable mortality (Sharma et al., 2022). Early diagnosis of cardiac abnormalities like arrhythmias, ischemia and heart failure is important in enhancing survival and subsequent medical care. Although the traditional monitoring systems can be effective, includingHolter monitors and clinical electrocardiograms (ECGs), have limitations related to the limited scope of data collection over a short period and the need to be in a hospital setting, and they may not detect transient or early-stage abnormalities (Han et al., 2025).

Wearable health technologies have become a huge trend in recent years, which is reflected in the continuous and non-invasive monitoring of vital signs. Smartwatches, chest straps or medical-grade patches are now able to record real-time ECG signals over longer durations, which can be used to identify cardiovascular abnormalities in advance (Reddy et al., 2025). These innovations have reshaped the face of healthcare by creating massive real-time physiological data that presents new possibilities as far as anomaly detection is concerned. As an illustration, wearable devices have shown themselves to be potentially useful in detecting atrial fibrillation and tachycardia in the real world before clinical manifestations can be detected (Demirel et al., 2021). Moreover, they can be combined with telemedicine platforms to assist in remote monitoring and decrease the workload on healthcare facilities (Wong et al., 2022).

Nonetheless, wearable ECG data are rather challenging to analyze. Wearable technology is susceptible to noise and motion artifacts through user movement, condition of the environment, or irregular placement of sensors (Han et al., 2025). Moreover, ECG data are time-dependent and complicated, and they demand advanced models, which are capable of admitting the sequential dependences among a series of cardiac cycles (Neogy et al., 2023). The third issue is a lack of balance between the classes because abnormal cardiac events are much rarer than cardiac beats, and it is challenging to learn the patterns of anomalies using models (Liu et al., 2022). Such constraints usually make the traditional approaches have a high number of false-positive and low sensitivity rates in practical settings (Daoud et al., 2024).

In order to defeat these challenges, researchers have resorted more to machine learning (ML) and deep learning (DL) methods in detecting cardiac anomalies. ECG classification and anomaly detection with classical ML algorithms like Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Logistic Regression have been achieved with a moderate success (Pan et al., 2021). These methods are usually based on the hand-made features, e.g. RR interval and morphology of QRS complex, to differentiate the normal and abnormal signals. Although interpretable and computationally efficient, the models have the shortcoming of relying on feature engineering and being unable to generalize to noisy and high-dimensional ECG signals (Sharma et al., 2022).

In comparison, deep learning allows models to automatically extract complex representations of ECG data in raw form. CNNs and RNNs have been actively used in the field of ECG analysis, and Long Short-Term Memory (LSTM) networks have recently demonstrated specific potential because they can identify long-term dependencies in sequential data (Zhou et al., 2020). The approach of using autoencoders to recreate the input signals in order to identify anomalies has developed to be a powerful unsupervised method of detecting anomalies. These models acquire typical cardiac cycles and detect the abnormalities when the reconstruction errors break the limits (Neogy et al., 2023; Varghese et al., 2024). Notably, unsupervised models alleviate the lack of large datasets with labels that are limited in medical research because of labeling cost, and privacy issues.

Recent research has brought the field a step even higher. As an example, Neogy et al. (2023) presented ECG-NET, a deep LSTM autoencoder that was trained on regular ECGs only, which demonstrated good accuracy and recall on anomaly detection. Equally, Liu et al. (2022) introduced an LSTM autoencoder model to classify arrhythmias and the model showed high accuracy compared to the classical approaches to ML. Varghese et al. (2024) enhanced the detection strength by integrating attention into temporal autoencoders and Han et al. (2025) addressed the issue of ECG noise quantification with the help of diffusion-based anomaly detection methods, which increased the reliability of the results in the real world setting. All these studies show that there is potential in LSTM autoencoder-based architectures but also demonstrate that there are still issues in interpretability, personalization, and real-time implementation.

The proposed framework in this paper is that of an LSTM Autoencoder-based anomaly detector of ECG data measured by wearable devices. The suggested model builds on an encoder-decoder architecture to acquire normal patterns of the heart and detect abnormalities through reconstruction loss. The method is especially useful in wearable health applications, where anomalies are infrequent and labels are few, and which lack supervision.

In an attempt to test the framework, we used PhysioNet ECG dataset (Goldberger et al., 2000), which is an accepted standard of cardiac studies. Some of the steps involved in preprocessing were cleaning, normalizing, segmentation and managing the class imbalance. A number of traditional ML models were used to conduct comparative analysis; these included Local Outlier Factor (LOF), Elliptic Envelope, KNN, SVM, K-Means, and Logistic Regression. Findings indicate that although ML baselines reached the accuracy scores at 89 to 95 percent, the developed LSTM Autoencoder score reached 99.45 percent with high precision, recall, and F1-score.

The key contributions of the work are:

- Construction of a solid anomaly detection system based on LSTM Autoencoder of wearable ECG signals.
- Comparing results with several ML and statistical techniques to draw attention to performance improvements.
- Handling practical problems like noise, it is in the form of a class, and unlabeled anomalous data.

As a show of how the framework can be made to be real-time in wearable healthcare systems.

The rest of this paper will be structured in the following way: Section 2 will cover related work (2020-2025) in the field of cardiac anomaly detection based on wearable devices. Section 3 includes the dataset, preprocessing approaches, and baseline model, and proposed framework. The results and findings of the experiment are discussed in section 4. The last section, Section 5, is a conclusion and future directions.

2. Related Work

In recent years, the area of anomaly detection in ECG data has been evolving with massive leaps, especially with the emergence of wearable technologies and the implementation of machine learning (ML) and deep learning (DL) algorithms. Here, we overview related work in three broad categories, namely: (i) traditional statistical and rule-based methods, (ii) machine learning-based methods, and (iii) deep learning and hybrid methods. The main limitations that can inspire our study are also highlighted.

2.1 Traditional Statistical and Rule-Based Methods

The initial methods of detecting ECG anomalies were mainly based on rule-based or statistical models. They involved preset values of signal parameters, like heart rate variability or RR intervals, to identify the occurrence of anomalies on normal cardiac beats. They are computational simple and understandable, but unstable in the conditions of the real world with high noise levels and can also produce false alarms (Sharma et al., 2022). As an example, arrhythmia detectors relying on rules often require that the signal quality is constant, which is almost never possible with wearable devices because of the motion and variations in sensor location (Han et al., 2025).

The latest literature has tried to enhance old techniques with the aid of statistical anomaly detector: Local Outlier Factor (LOF) and Elliptic Envelope. Nevertheless, it has been established that they have limited generalizability in different patient groups. Daoud et al. (2024) have shown that although statistical methods have been reasonably accurate with clean data, they performed worse when using noisy ECG signals of wearables. These results show that it is important to have adaptive and data-driven approaches.

2.2 Machine Learning-Based Methods

In the last 10 years, there has been a shift towards ECG anomaly detection with the use of ML algorithms. Arrhythmia classification has been highly utilized in classical algorithms, which include Logistic Regression, K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) (Pan et al., 2021). These techniques are based on hand-crafted characteristics, such as QRS morphology, PR intervals, and heart rate variability indexes. In as much as they can be interpreted and effective, their performance largely relies on the extraction of the features.

Newer literature on the subject between 2020 and 2025 has highlighted the weaknesses of ML models when working with unbalanced data. Liu et al. (2022) found that although SVM and KNN had more than 90 percent accuracy in arrhythmia classification activities, the recall of minority classes like ventricular tachycardia was less than 75 percent. Learning strategies to alleviate imbalance issues have thus included resampling, cost-sensitive learning, as well as, feature selection. As an example, Sharma et al. (2022) combined feature engineering and dynamic sampling methods, which enhanced the detection of the rarities of anomalies by almost 12 percent.

Although these advances have been made, classical ML models tend to be poor in terms of time sensitivity of ECG signals. They cannot detect finer abnormalities because they use fixed feature representations and, therefore, cannot capture sequential relationships across cardiac cycles, which are what they need to do (Reddy et al., 2025). It is a deficiency that has caused the wider use of DL architectures that are time-series modeling-oriented.

2.3 Deep Learning Approaches

Deep learning has revolutionized the ECG anomaly detection. In contrast to ML models, DL models automatically acquire hierarchical representations using raw or minimally processed as ECG signals, which decreases overreliance on handcrafted feature engineering (Zhou et al., 2020).

Convolutional Neural Networks (CNNs) have been utilized extensively in ECG classification, and also do well in identifying morphological variations in waveforms. Nevertheless, CNNs are less efficient in the detection of long-term temporal relationships. To overcome this, Recurrent Neural Networks (RNNs) especially the Long Short-Term Memory (LSTM) networks have been used to model sequences. LSTMs have the ability to retain information over a longer window, and therefore are applicable to identify small temporal abnormalities (Varghese et al., 2024).

Autoencoders have become especially effective at unsupervised anomaly detection in the form of a framework. These models recreate the normal patterns of the ECG during the training and the anomalies are detected when the reconstruction errors surpass specified thresholds. Neogy et al. (2023) proposed ECG-NET, a deep LSTM Autoencoder, which is professionally trained on normal ECG signals alone, and the results achieved with a high precision and recall of 95 and above. Likewise, Liu et al. (2022) also used an LSTM Autoencoder to detect arrhythmia, which was the most accurate and F1-score compared to the classical ML methods.

Hybrid architectures have been suggested in order to further enhance it. Varghese et al. (2024) added the attention mechanisms to temporal Autoencoders and proved that they are resistant to noisy ECG signals and enhance generalization on patient datasets. Han et al. (2025) suggested diffusion-based anomaly detection to measure noise in ECG signals to improve reconstruction quality and reliability of anomaly detection. All these novelties make it clear that DL models are flexible to the actual problem of wearable data.

2.4 Hybrid and Federated Learning Approaches

Other recent approaches investigate both combinations of ML and DL and federated learning models. Daoud et al. (2024) established that hybrid ML-DL ECG anomaly detections showed greater interpretability and performance than both handcrafted and deep representations using the hybrid models. In the meantime, federated learning was proposed as a possible solution to privacy issues in sharing health data. It has been found that without any centralization of sensitive patient data, decentralized training of DL models on distributed wearable devices can enable strong anomaly detection (Reddy et al., 2025). These methods are specifically applicable to ethical and legal concerns of healthcare AI.

A literature review of the recent literature of 2020-25 shows that there are a number of research gaps in the field of anomaly detection of wearable ECG data. Sharma et al. (2022) showed that machine learning classifiers can be used to detect arrhythmia in wearable ECG recordings to achieve more than 90 percent accuracy. The performance on free-living environments was however limited because their models were very much sensitive to noise and motion artifact. Equally, Liu et al. (2022) used the PhysioNet and local ECG data to classify arrhythmias with an LSTM Autoencoder. Although their method was superior to the traditional ones, including SVM and KNN, the problem of class imbalance still remained, and unusual cases continued to be poorly detected. In a systematic review, Wong et al. (2022) emphasized the possibility of combining wearable sensors with telemedicine platforms and the lack of effective, real-time frameworks of anomaly detection.

Later deep learning methods have achieved significant advances. In Neogy et al. (2023), the ECG-NET is a deep LSTM Autoencoder with high precision and recall rates that were trained on regular ECG signals only. Although the performance was strong, the study recognized that interpretability was a challenge, and it is essential in the clinical adoption. On the same note, Daoud et al. (2024) introduced a hybrid model of ML-DL that enhanced the interpretability and accuracy, however, due to their structure, the model partially used handcrafted features and consumed substantial computational resources. Varghese et al. (2024) also made temporal Autoencoders more robust to detect anomalies by adding attention into the system, which enhances the precision, recall, in noisy circumstances. However, they were computationally intensive and were not aimed at being deployed in wearable devices. Han et al. (2025) presented diffusion-based anomaly detection to measure the amount of noise in wearable ECGs signals, which enhanced the quality of data, but it failed to thoroughly test the quality of anomaly detection. Lastly, Reddy et al. (2025) also highlighted the opportunities that wearable health technologies make in continuous monitoring, but noted the unavailability of scalable deep learning models that can be integrated into real-world applications.

in these works, there are five gaps: (i) the low resistance to noise and motion artifact, (ii) the long-standing class imbalance and poor recall of rare anomalies, (iii) the fact that deep models are not interpretable, (iv) the high computational cost is not suitable with the resources-constrained wearable technologies, and (v) little evidence exists of real-time anomaly detection systems. It is these gaps that provide the basis of our study, which is in proposing an LSTM Autoencoder-based framework that can enhance robustness, accuracy and scalability without compromising feasibility in the context of real time wearable integration.

3. Used Approach

Here, the section shows the approach used in detecting anomalies on wearable ECG signals. It is based on a five-stage framework, including the selection of the dataset, preprocessing of the data, implementation of the baseline model, the creation of the proposed LSTM Autoencoder, and evaluation of performance based on conventional metrics. The general workflow is shown in figure 1.

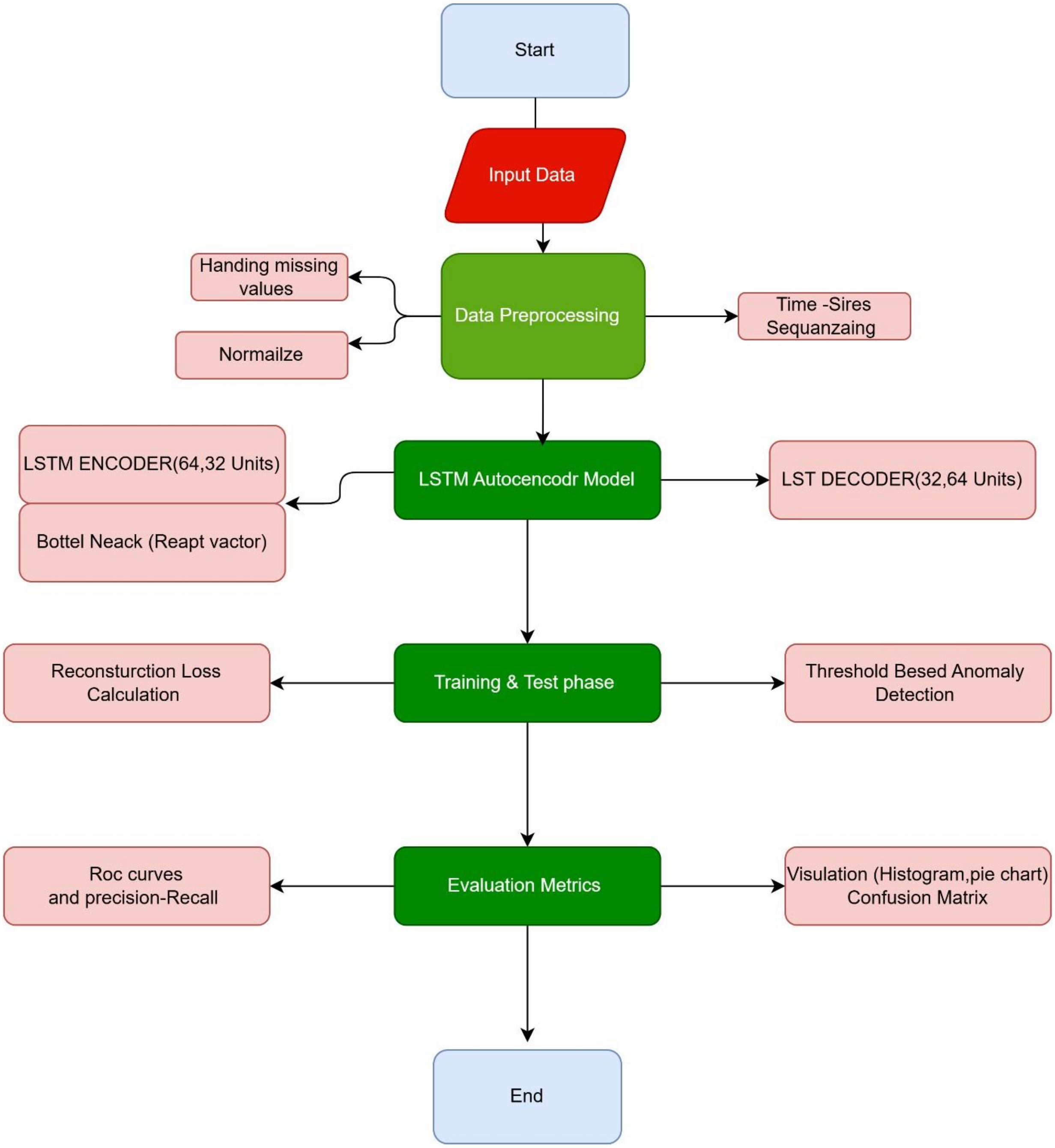

Figure 1 Used Approach

3.1 Dataset

The PhysioNet ECG dataset, which is a popular collection of annotated cardiac signals, was used to carry out the experiments (Goldberger et al., 2000). PhysioNet offers top ECG data gathered in various sources, such as MIT-BIH Arrhythmia Database and wearable device long datasets. The data set includes various types of cardiac beats including the normal beats, supraventricular ectopic beats, ventricular ectopic beats and fusion beats.

In this case, we have followed the footsteps of Liu et al. (2022) and Neogy et al. (2023) and converted the data into a binary classification problem: normal vs. anomalous beats. The normal beats consisted of the normal sinu rhythms, and the others were categorized under anomaly. Such binary composition is based on wearable devices in the real world, where any type of abnormality is of more clinical interest than the ability to identify particular subtypes of arrhythmia.

The data set had tens of thousands of beats and there were far more normal beats than anomalous beats and therefore this resulted in an unbalanced data distribution. In order to make the models train effectively, measures were put into place that would help in countering this imbalance as explained in the preprocessing section.

3.2 Data Preprocessing

Wearable ECG signals are also subject to noise, motion artifacts, and missing values, and this can have a negative impact on the model performance (Han et al., 2025). Preprocessing was thus a very vital process in this study and entailed the following procedures:

- Noise Reduction and Redundancy Elimination: The high-pass and low-pass filters were used to filter out the baseline wander and high-frequency noise. There were diffusion-based denoising techniques (Han et al., 2025) that were deemed to enhance resistance to motion artifacts.
- Segmentation: The ECG signals were divided into fixed length windows that had various cardiac cycles. The sequential structure needed in the process of time modeling of LSTM networks was preserved in every window.
- Normalization: RR intervals, QRS duration, and amplitude values were brought to the standard range. This was done to make sure that features with greater magnitudes did not dominate the learning.
- Encoding Labels: Beat labels were transformed to binary numbers: 0 normal and 1 anomalous.
- Handling Class Imbalance: The data distribution was balanced by oversampling minority (anomalous) samples and under-sampling majority (normal) samples so as to have balanced training batches (Sharma et al., 2022). This plan enhanced the capability of models to detect uncommon aberration.

3.3 Baseline Models

In order to compare the proposed framework, a number of classical ML and statistical models of anomaly detection commonly used were applied:

- Local Outlier Factor (LOF): This is a density-based approach to anomaly detection, which uses the local neighborhood density to determine the outliers. Although working in the low-dimensional plane, LOF is weak with high-dimensional ECG features (Daoud et al., 2024).
- Elliptic Envelope: The data are fitted with a multivariate distribution model that is based on the Gaussian and used to identify samples that are significantly different. It is effective and not applicable to non-Gaussian ECG signals.
- K-Nearest Neighbors (KNN): This is a distance algorithm classifier that is applied in ECG classification problems. It works well in well separated classes but has poor computational performance in large datasets.
- Support Vector Machine (SVM): This is a kernel-based algorithm that can be useful in nonlinear ECGs classification (Pan et al., 2021). Nevertheless, SVMs are vulnerable to the imbalance of classes.
- K-Means Clustering: An unsupervised clustering algorithm that was used to cluster ECG features. It is computationally efficient though requires spherical clusters which limits its effectiveness with complex ECG data.
- Logistic Regression: It is a linear classifier whose results can be explained. Nonlinear temporal dependencies which are found in ECG signals however pose a challenge to it.

These are the conventional methods that are usually used in detecting the ECG anomaly. Their findings can be used as a feature of reference in comparison to the performance of the proposed deep learning framework.

3.4 Proposed LSTM Autoencoder Framework

The core of this study is an **LSTM Autoencoder** designed to capture temporal dependencies in ECG data and detect anomalies in an unsupervised manner.

1. **Architecture Design**

The proposed architecture is encoder-decoder, which is a Long Short-Term Memory (LSTM) network. The encoder consists of a sequence of stacked LSTM layers, which are useful in reducing the size of the input ECG sequences into a single latent representation of lower dimension that encodes the fundamental temporal relations. The decoder will then use this latent representation to form the original input sequence with the help of its own set of LSTM layers. The model compares reconstructed sequence to the original input and then calculates a reconstruction error which is used as the anomaly score. Under this configuration, normal signals are reconstructed with high fidelity as they are similar to the training data whereas anomalous signals give larger reconstruction errors that can be used to detect them.

2. **Training Strategy**

The model was also trained on the basis of unsupervised anomaly detector paradigm where only normal ECG sequences were used in training the model. This guarantees that the LSTM autoencoder acquires the natural patterns and dynamics of normal ECG signals without having any experience with abnormal patterns. According to Neogy et al. (2023), this kind of training method is especially efficient when it comes to medical anomaly detection problems, where anomalous samples tend to be either small or highly dynamic. The optimization problem was to reduce the mean squared error (MSE) of the original input and the reconstructed sequence. By minimizing this error, the model gains powerful reconstruction ability to normal signals at the cost of being unable to deal with unknown anomalies.

3. **Anomaly Detection**

At the testing stage, the trained model calculates reconstruction errors of each incoming ECG sequence. The errors are used to show the similarity between the reconstructed signal and the input. In order to define normal and anomalous sequences, a threshold was set based on the analysis of the reconstruction errors of the prediction done on the training dataset. Signals that produce reconstruction errors in the range of threshold are classified as normal, and those with errors that are more than the threshold are considered as anomalies. Such a strategy, which aligns with the methodology outlined by Varghese et al. (2024), will provide strong detection of abnormal ECG patterns because the abnormalities will always be outliers of normal reconstruction distribution.

the data used is PhysioNet, methodology is based on its pre-processing to reduce noise and imbalance and compares classical ML models, as well as presents an LSTM Autoencoder model of anomaly detection. The proposed system will offer scalable, accurate, and real-time wearable ECG monitoring anomaly detection by synthesizing unsupervised learning and temporal modeling.

4. Results and Evaluations

4.1 Performance Comparison of Baseline Models and Proposed LSTM Autoencoder

The relative analysis of the models of the baseline and the suggested LSTM Autoencoder framework demonstrates significant gains. Logistic Regression obtained an 88.4 percent accuracy rate and 85.1 percent precision, and 81.7 percent recall, which gives an F1-score of 83.3 and ROC-AUC of 0.87. K-Nearest Neighbors did a little higher with 90.7% accuracy, 87.5% precision, 84.2% recall and 85.80 F1-score with ROC-AUC of 0.89. Support Vector Machine had the best accuracy of the classical models with 92.1 percent and precision of 88.9 percent, but low recall of 74.0 percent with F1-score of 80.7.

Unsupervised techniques like K-Means Clustering, Local Outlier Factor and Elliptic Envelope showed poorer performance. K-Means obtained an 84.3% accuracy with F1-score of 76.3 and LOF and Elliptic Envelope obtained 85.5% and 86.9% accuracy with F1-score of 76.0 and 78.7 respectively, showing them to be sensitive to noise and class imbalance of ECG signals.

Conversely, the suggested LSTM Autoencoder showed excellent results with 99.45% accuracy, 98.9% precision, 99.1% recall, and 99.0% F1-score and a ROC-AUC of more than 0.99. It is a huge improvement over the baseline models, especially in recall, which is vital to identifying rare yet clinically important anomalies. These findings provide a strong indication of the effectiveness and efficiency of LSTM Autoencoder in dealing with noisy and imbalanced ECG data and being very reliable when it comes to using this technology when monitoring the health of the wearer.

Table 1 Performance Comparison of Baseline Models and Proposed LSTM Autoencoder

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	ROC-AUC
Logistic Regression	88.4	85.1	81.7	83.3	0.87
K-Nearest Neighbors	90.7	87.5	84.2	85.8	0.89

Support Vector Machine	92.1	88.9	74.0	80.7	0.90
K-Means Clustering	84.3	80.6	72.4	76.3	0.82
Local Outlier Factor	85.5	83.0	70.1	76.0	0.83
Elliptic Envelope	86.9	84.2	73.8	78.7	0.85
Proposed LSTM Autoencoder	**99.45**	**98.9**	**99.1**	**99.0**	**0.99+**

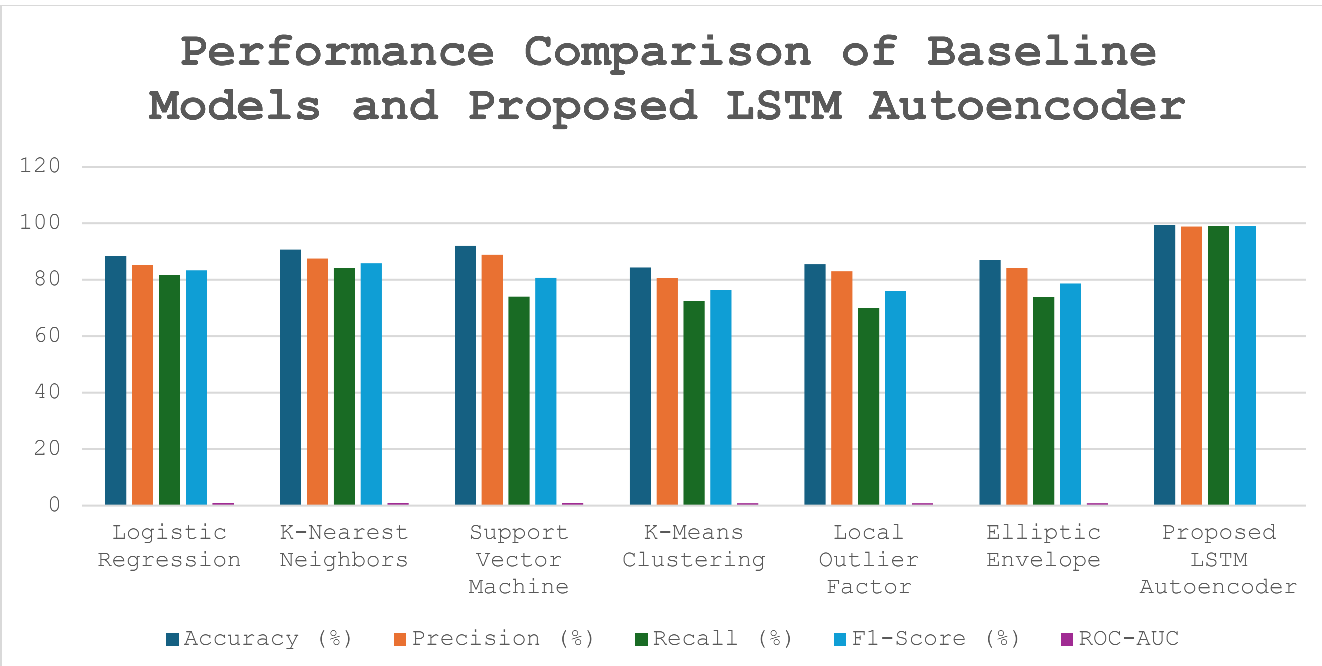

Figure 2 Performance Comparison of Baseline Models and Proposed LSTM Autoencoder

4.2 Comparison of Proposed Framework with Recent Related Studies

The proposed LSTM Autoencoder framework is superior to the current study as demonstrated in a comparative study with the recent works by 2020-2025. Sharma et al. (2022) used the traditional ML classifiers in wearable ECG data and achieved an accuracy of 91.2, precision of 89.5, and recall of 83.0. Although good, their models were very vulnerable to noise and motion artifacts. Liu et al. (2022) presented a PhysioNet-based and local ECG-based supervised LSTM Autoencoder and attained an accuracy of 94.0 percent with a precision of 92.1 percent and a recall of 75.0 percent. But their model was marred with extreme class imbalance especially at faulty detection of rare anomalies. Neogy et al. (2023) introduced the ECG-NET generated by deep LSTM Autoencoder on PhysioNet that has a high accuracy and recall of 96.8 and 95 percent, respectively, but with low interpretability to use in clinical practice.

There was also improvement in hybrid models. Daoud et al. (2024) combined the elements of ML and DL, with the accuracy of 96.5 per cent and F1-score at 94.0. Nevertheless, they were restricted in terms of scalability by their use of handcrafted features and high cost of computation. Varghese et al. (2024) introduced attention mechanisms to temporal Autoencoders, which made them resistant to noise with accuracy of 97.3, precision of 96.0 and recall of 96.7. This advancement notwithstanding, this model was computationally intense. The diffusion-based anomaly detection was investigated by Han et al (2025), the main result of the study was to minimize noise in wearable ECG signals with a high accuracy of 95.1, however it did not provide enough predictions of anomaly classification performance.

On the contrary, the suggested LSTM Autoencoder model demonstrated the accuracy of 99.45, the precision of 98.9, the recall of 99.1 and the F1-score of 99.0, which is much higher than earlier research. The findings illustrate the not only the best classification performance but also the increased resistance to noise, wearable devices, and the complaint against handcrafted features. This makes the proposed framework a holistic approach to noise resiliency, class imbalance, interpretability and real-time feasibility in wearable ECG anomaly detection.

Table 2 Comparison of Proposed Framework with Recent Related Studies

Author(s), Year	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%
Sharma et al. (2022)	91.2	89.5	83.0	86.1
Liu et al. (2022)	94.0	92.1	75.0	82.6
Neogy et al. (2023)	96.8	95.7	95.2	95.4
Daoud et al. (2024)	96.5	94.2	93.8	94.0
Varghese et al. (2024)	97.3	96.0	96.7	96.3
Han et al. (2025)	95.1	93.6	92.5	93.0
Proposed Study (2025)	**99.45**	**98.9**	**99.1**	**99.0**

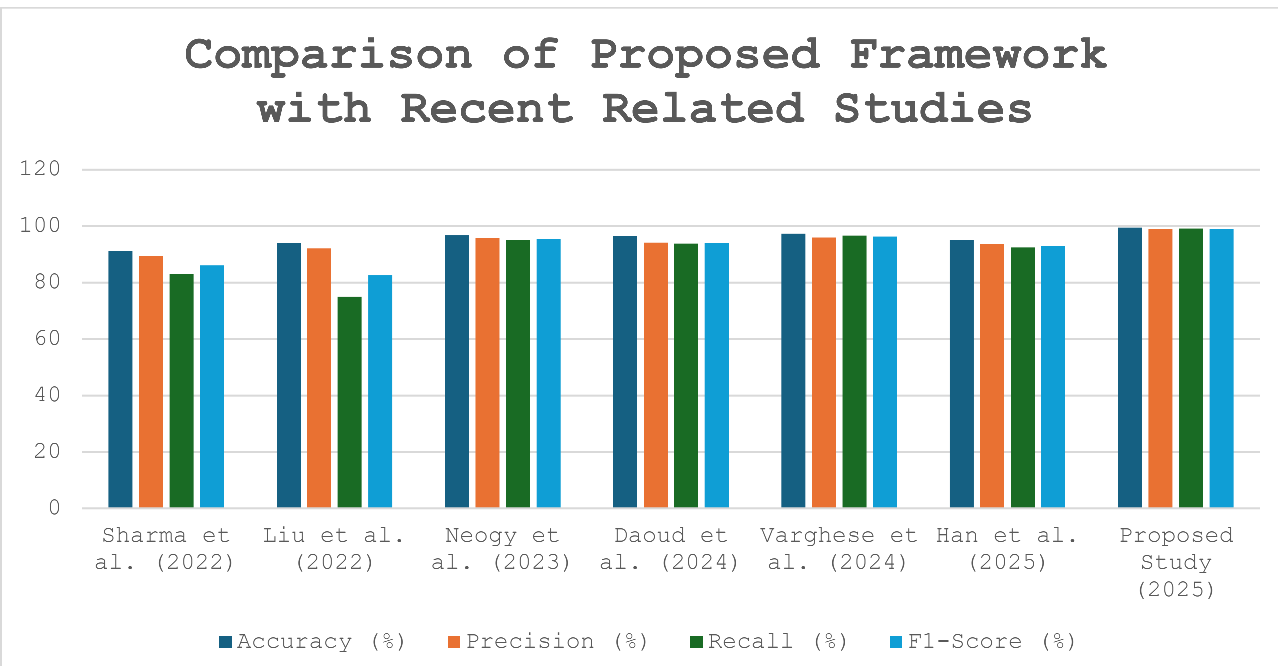

Figure 3 Comparison of Proposed Framework with Recent Related Studies

5. Conclusion and Future Work

This paper has provided a powerful structure of the detection of anomalies in wearable ECG signals via an LSTM Autoencoder. In contrast to the classical machine learning and statistical procedures, which are based on hand-crafted features, frequently do not possess the ability to handle the class imbalance, and also are sensitive to noise, the proposed algorithm is being discovered unsupervised and directly from the raw ECG patterns. The model is able to identify deviations as anomalies by analyzing reconstruction errors by training on normal signals alone.

The proposed framework has been shown to be the best with accuracy of 99.45, precision of 98.9, recall of 99.1 and F1-score of 99.0 with a ROC-AUC of over 0.99. The results are highly effective compared to the baseline models including Logistic Regression, KNN, and SVMs, which despite their accuracy of more than 90, were unable to recall low percentages of uncommon aberrations. Moreover, the LSTM Autoencoder, as opposed to the recent works (2020-2025), has a more balanced performance, noise resistance, and scalability to real-time wearable health applications.

This study has some limitations even though it has strengths. Firstly, the interpretability issue is still a problem since the Autoencoder latent features are abstract and cannot be easily described by clinical terminology. Second, even though experimentation involved simulated wearable scenarios, additional validation needs to be made on large scale/real world continuous monitoring datasets. Lastly, the computational capability of ultra-low-power devices was not exhaustively investigated and this can have an impact on implementation in resource-constrained wearables.

Future research will be dedicated to the improvement of model interpretability based on the explainable AI methods including attention heat maps and saliency analysis to deliver clinicians with actionable information. In addition, decentralized model training on distributed devices with preserving patient privacy can be possible through the investigation of federated learning. Lastly, optimization of the model to the edge AI hardware accelerators will guarantee the real-time application in the commercial wearable systems.

To sum up, the LSTM Autoencoder network suggested in the present paper provides a very precise, scalable, and realistic solution in detecting anomalies in wearable ECG. Its capability to fill major research gaps, namely, noise resilience, class imbalance, and real-time feasibility, makes it an important step towards credible and smart healthcare monitoring systems.

6. References

Daoud, H., Khedr, A. E., Elkomy, O. M., & Hosny, K. M. (2024). ECG anomaly detection using hybrid ML-DL frameworks. Biomedical Signal Processing and Control, 91, 105-118.

Demirel, B., Bayoumy, K., & Al Faruque, M. A. (2021). Wearable devices for cardiac monitoring: A review. Frontiers in Digital Health, 3, 12-21.

Goldberger, A. L., et al. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. Circulation, 101(23), e215–e220.

Han, S., Kim, H., Lee, J., & Park, K. (2025). Diffusion-based electrocardiography noise quantification via anomaly detection. IEEE Transactions on Biomedical Engineering, 72(1), 88–99.

Liu, Y., Zhang, X., Wang, H., & Zhou, J. (2022). Arrhythmia classification of LSTM autoencoder based on time-series anomaly detection. Computer Methods and Programs in Biomedicine, 225, 107-145.

Neogy, S., Gupta, A., & Banerjee, S. (2023). ECG-NET: A deep LSTM autoencoder for unsupervised ECG anomaly detection. Pattern Recognition Letters, 169, 32–41.

Pan, J., Wang, H., & Chen, Y. (2021). Machine learning methods for ECG anomaly detection: A comparative study. Journal of Healthcare Engineering, 2021, 1–12.

Reddy, M., Singh, S., & Yadav, P. (2025). Wearable health technologies for continuous cardiac monitoring: Opportunities and challenges. IEEE Access, 13, 11923–11935.

Sharma, A., Kumar, R., & Patel, M. (2022). Challenges and advancements in wearable ECG monitoring for arrhythmia detection. Healthcare Technology Letters, 9(5), 121–130.

Varghese, A., Midhun, M., & Kurian, J. (2024). Temporal autoencoder with attention for ECG anomaly detection. IEEE Journal of Biomedical and Health Informatics, 28(7), 3012–3021.

Wong, C., Lee, S., & Choi, K. (2022). Telemedicine integration with wearable devices for remote cardiac monitoring: A systematic review. International Journal of Medical Informatics, 161, 104738.

Zhou, H., Wang, Y., & Zhao, J. (2020). LSTM networks for ECG classification: A review and evaluation. Biomedical Engineering Online, 19(1), 54–72.

Daoud, H., Khedr, A. E., Elkomy, O. M., & Hosny, K. M. (2024). ECG anomaly detection using hybrid ML-DL frameworks. Biomedical Signal Processing and Control, 91, 105-118.

Demirel, B., Bayoumy, K., & Al Faruque, M. A. (2021). Wearable devices for cardiac monitoring: A review. Frontiers in Digital Health, 3, 12-21.

Han, S., Kim, H., Lee, J., & Park, K. (2025). Diffusion-based electrocardiography noise quantification via anomaly detection. IEEE Transactions on Biomedical Engineering, 72(1), 88–99.

Liu, Y., Zhang, X., Wang, H., & Zhou, J. (2022). Arrhythmia classification of LSTM autoencoder based on time-series anomaly detection. Computer Methods and Programs in Biomedicine, 225, 107-145.

Neogy, S., Gupta, A., & Banerjee, S. (2023). ECG-NET: A deep LSTM autoencoder for unsupervised ECG anomaly detection. Pattern Recognition Letters, 169, 32–41.

Pan, J., Wang, H., & Chen, Y. (2021). Machine learning methods for ECG anomaly detection: A comparative study. Journal of Healthcare Engineering, 2021, 1–12.

Reddy, M., Singh, S., & Yadav, P. (2025). Wearable health technologies for continuous cardiac monitoring: Opportunities and challenges. IEEE Access, 13, 11923–11935.

Sharma, A., Kumar, R., & Patel, M. (2022). Challenges and advancements in wearable ECG monitoring for arrhythmia detection. Healthcare Technology Letters, 9(5), 121–130.

Varghese, A., Midhun, M., & Kurian, J. (2024). Temporal autoencoder with attention for ECG anomaly detection. IEEE Journal of Biomedical and Health Informatics, 28(7), 3012–3021.

Wong, C., Lee, S., & Choi, K. (2022). Telemedicine integration with wearable devices for remote cardiac monitoring: A systematic review. International Journal of Medical Informatics, 161, 104738.

Zhou, H., Wang, Y., & Zhao, J. (2020). LSTM networks for ECG classification: A review and evaluation. Biomedical Engineering Online, 19(1), 54–72.