Journal of Theoretical and Applied Information Technology
15th March 2026. Vol.104. No.5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INTEGRATING ADVERSARIAL AUTOENCODERS WITH GATED RECURRENT UNITS TO IDENTIFY ANOMALIES IN ECG SIGNALS

SHAIK JANBHASHA[1], VENKATESWARLU SUNKARI[2], DIPANWITA DEBNATH[3], VENKAT RAO PASUPULETI[4], AKETI NARESH[5], LAKSHMANARAO TALAPAKULA[6], JUTU GOPAIAH[7], ROHTIH BALA JASWANTH B[8]

[1]Associate professor, CVR College of Engineering, Ibrahimpatnam, Hyderabad, Telangana, Department of Computer Science and Engineering (Data Science), India.
[2]Assistant Professor, University of Nizwa, Birkat Al Mouz, Nizwa 616, Department of Electrical and Computer Engineering, Oman,
[3]Assistant Professor, B V Raju Institute of Technology, Narsapur, Medak, Telangana, Department of Computer Science and Engineering (AI&ML), India.
[4]Associate Professor, Lakireddy Bal Reddy College of Engineering, Mylavaram, Andhra Pradesh, Department of ECE, India.
[5]Professor, Sri Sai Institute of Technology and Sciences, Rayachoty, Department of Computer Science and Engineering, India.
[6]Assistant Professor, Sasi Institute of Technology and Engineering, Tadepalligudem, Andhra Pradesh, Department of Computer Science and Engineering, India.
[7]Assistant Professor, Koneru Lakshmaiah Educational Foundation, Guntur, Andhra Pradesh, Department of Computer Science and Engineering, India.
[8]Associate Professor, Seshadri Rao Gudlavalleru Engineering College, Gudlavalleru, Krishna District, Andhra Pradesh, Department of Internet of Things, India.
Email: [1]afreen.jbasha@gmail.com, [2]v.sunkari@unizwa.edu.om, [3]dipanwita.d@bvrit.ac.in, [4]venkat2k15@gmail.com, [5]pandu5188@gmail.com, [6]lakshman@sasi.ac.in, [7]gopi.jutu32@gmail.com, [8]jaswanthiot@gmail.com

ABSTRACT

Cardiovascular diseases remain a leading global health concern, necessitating advanced methods for early detection of cardiac anomalies. This study proposes a hybrid model combining Adversarial Autoencoders (AAE) and Gated Recurrent Units (GRU) to improve anomaly detection in Electrocardiogram (ECG) signals. The AAE module extracts robust latent representations of normal ECG patterns, while the GRU captures temporal dependencies within the signal. Experimental results demonstrate superior performance compared to existing methods, achieving 98.6% accuracy, 97.7% recall, 98.2% F1-score, 98.8% precision, and 99.1% AUC-ROC. The proposed framework reduces false positives and enhances diagnostic reliability, offering a promising tool for automated cardiac monitoring.
Keywords: *ECG, Anomaly Detection, Adversarial Autoencoder, GRU, Deep Learning*

1. INTRODUCTION

Electrocardiogram (ECG) signals are essential for diagnosing and monitoring a wide array of cardiac abnormalities, including arrhythmias, myocardial infarctions (heart attacks), and atrial fibrillation. These conditions remain a leading cause of death worldwide, contributing to over 17 million deaths annually. Early detection of these cardiac anomalies is crucial for improving patient outcomes, as timely medical intervention can significantly reduce the risk of life-threatening events. However, traditional ECG analysis relies on manual interpretation by clinicians, a process that is not only time-consuming but also prone to human error due to the complex and noisy nature of ECG data. As a result, automated ECG analysis systems are being increasingly sought after to augment the diagnostic capabilities of healthcare professionals and enable continuous, real-time monitoring of patients [1].

The Figure 1 illustrates a typical ECG (electrocardiogram) waveform, representing the electrical activity of the heart during one cardiac cycle. The P wave signifies atrial depolarization, leading to atrial contraction. The PR segment reflects the delay at the AV node, allowing blood to flow from the atria to the ventricles. The QRS complex indicates rapid ventricular depolarization, triggering ventricular contraction.

Journal of Theoretical and Applied Information Technology
15th March 2026. Vol.104. No.5

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

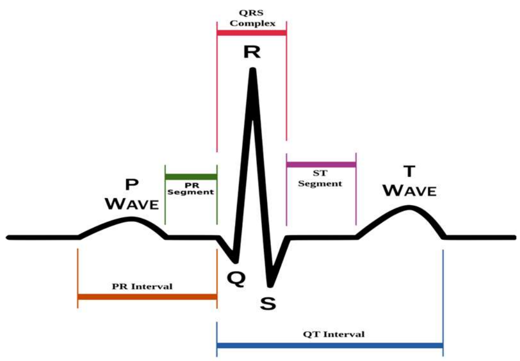

Figure 1: ECG Signal [5]

Following this, the ST segment represents the early phase of ventricular repolarization. The T wave shows the ventricles returning to their resting state. The PR interval spans from the beginning of the P wave to the start of the QRS complex, and the QT interval extends from the beginning of the QRS to the end of the T wave, representing total ventricular activity. This waveform helps diagnose various cardiac conditions, including arrhythmias and myocardial infarction [16].

Traditional ECG analysis methods, including rule-based systems and shallow machine learning models like SVMs and decision trees, face limitations in detecting subtle or complex abnormalities. These approaches depend heavily on handcrafted features and domain expertise, reducing their generalizability across diverse patient populations or novel arrhythmias [21]. They also struggle to capture temporal dependencies in ECG signals, often missing rare or evolving cardiac events and increasing the risk of misdiagnosis or false alarms [2].

Deep learning overcomes traditional ECG analysis limitations by learning features directly from raw data. Models like CNNs and GRUs effectively capture spatial and temporal patterns in ECG signals [22]. CNNs extract local features, while GRUs model sequential dependencies for accurate classification. However, real-world deployment is challenged by class imbalance and signal noise, which can hinder performance—especially in unsupervised settings with limited labelled data [3].

This study presents a hybrid model combining an Adversarial Autoencoder (AAE) and a Gated Recurrent Unit (GRU) to enhance ECG anomaly detection. The AAE learns the distribution of normal signals in an unsupervised manner, enabling it to detect rare or evolving anomalies [20]. Adversarial training improves robustness by generating features that mimic normal data. The GRU captures long-term temporal patterns, allowing for the identification of subtle changes in cardiac activity [4]. This approach addresses challenges like class imbalance and signal noise in real-world ECG analysis.

By combining AAE and GRU, the proposed hybrid model achieves high anomaly detection accuracy while remaining robust to noise and minimizing false positives. This is critical for real-time monitoring systems where timely detection can improve patient outcomes. The AAE enables unsupervised learning, enhancing adaptability to new patient data and reducing reliance on large labeled datasets [23]. Our results show that the model outperforms traditional methods and offers a scalable solution for real-world ECG analysis [5].

Cardiovascular diseases (CVDs) constitute the major cause of mortality in the world and early and precise identification of ECG abnormalities is critical to treating this condition. The manual and rule-based analysis of ECG is prone to errors and noise, imbalance between classes, and time dependencies. Though deep learning algorithms like CNNs and RNNs are better, they need huge labeled datasets and cannot be used well in unsupervised tasks. Autoencoders and standalone recurrent models also do not provide sufficient structure to latent representations and represent long-term cardiac dynamics. In order to overcome these shortcomings, this paper provides a hybrid Adversarial Autoencoders-Gated Recurrent Unit (AAE-GRU) network to unsupervised ECG anomaly detection. The model adapts strong latent behaviour of normal ECG behaviour, lowering false positives and enhancing detection and generalisation of various patterns of ECG.

This paper will hypothesize that Adversarial Autoencoders combined with Gated Recurrent Units (AAE-GRU) are significantly better in detecting ECG anomalies by learning adversarial regularized latent representations of normal cardiac events. Another assumption is that the GRU-based temporal modeling boosts the detection of latent and dynamic abnormalities to the ECG by effectively learning long-term dependence in sequential signals. The analysis of reconstruction-error and latent-space discrimination combined is likely to decrease the false positives in the context of noisy and imbalanced conditions. In general, it is hypothesized that the AAE-GRU model will be more successful in comparison with the traditional deep learning models, such as CNN-LSTM, in regard to Accuracy,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

F1-score, and AUC-ROC, on regular ECG datasets[24].

2. LITERATURE REVIEW

The methods of ECG anomaly identification vary between traditional statistical models to the advanced deep learning tools, but these two classes have significant weaknesses. The classical approaches to machine learning including Support Vector Machines and Random Forests have moderate performance (approximately 85 percent accuracy) and are highly dependent on the hand crafted features and signal stationarity. As a result, they have trouble with non-stationary ECG signals, inter-patient variability and are limited in terms of their use in constant and real-world cardiac monitoring. Recurrent network models like LSTM enhance the temporal dependency modeling, and are found to be more accurate (up to 92%), but are supervised and prone to imbalance in the classes, which limits scalability in unlabeled clinical context. Equally, CNNs are good enough in identifying local morphological characteristics but not long-range temporal dependencies that are important to identify anomalies at rhythm level [17].

The Adversarial Autoencoders (AAEs) proved to be useful in medical anomaly detection through latent distribution regularization, enhancing the separation of normal and abnormal patterns. Simultaneously, GRUs are also effective in sequencing data with reduced case of vanishing gradient. Whereas recent combinations of autoencoders and recurrent networks have shown better performance (up to 94% AUC-ROC) these networks do not have adversarial regularization of the latent space, leading to worse representation robustness and increased false positive rates in rare or changing arrhythmias. This weakness shows that there is a major research gap in the collaboration of adversarial regularization with temporal modeling in the detection of ECG anomalies [18].

A number of studies based on CNN indicate high accuracy of classification when benchmark datasets are used. As an example, the 1D-CNN networks trained on the MIT-BIH Arrhythmia Database responded to greater than 91 percent of the cases with high specificity indicating fewer false alarms in a controlled setting [19]. Nevertheless, these models are mainly oriented to the supervised classification process and learning local features and, therefore, they are not as useful in identifying unseen or hidden anomalies. Equally, deep CNN models are better at rhythm classification than human cardiologists, thus highlighting their diagnostic capabilities, but are not as robust in long-term monitoring conditions due to their high instructions of large labeled datasets and continuous time modeling [7].

There is continual learning, which overcomes catastrophic forgetting and enhances flexibility over time, but adds extra model complexity and computational cost, which is difficult to implement in real-time healthcare systems [8]. Different patients populations are strongly associated with better generalization of models being trained with large-scale ECG datasets, as it was demonstrated in the examples of particular patients [9], [10]. However, dataset improvements are not the solution to inherent architectural weaknesses on the latent space structure, noise resilience and unsupervised anomaly recognition.

Hybrid and ensemble deep learning models, such as CNN-LSTM, all demonstrate high accuracy, sensitivity and specificity in various cardiac conditions. Such models are useful in capturing spatial and temporal ECG characteristics and are resistant to noise and variations in signal [11], [13], [14]. Although they perform well, the majority of hybrid architectures are still supervised and do not provide principled latent space regularization and, therefore, do not have the capacity to detect rare, changing, or never-before-seen anomalies. Transformer models also enhance the long-term dependency modeling, but with high computation and memory costs which may not be practical to use in long-term monitoring of ECG continuously [12].

The sequence-to-sequence models are more effective in inter-patient generalization and are very sensitive in the detection of critical arrhythmias like PVCs [15]. Nevertheless, their functions usually rely on large amounts of training data and a controlled experimental setting, which would be questioned regarding the practice of generalization in real clinical settings. On the whole, the literature indicates that deep learning models with excellent classes are often unaware of unsupervised adaptability, learned latent representation, and reduction of false positives. These loopholes inspire why a cohesive AAE-GRU architecture is required which combines adversarial latent space control with socially effective temporal modelling to convey powerful, scalable and clinically dependable ECG anomaly detection.

The reliability of the traditional and shallow methods is limited because noise, class imbalance, and limited labeled data make the correct detection of ECG anomalies difficult. Though deep learning algorithms like CNN-LSTM are more effective, they require a massive amount of labeled data and cannot be used in unsupervised cases. The current autoencoders as well as recurrent methods lack the simultaneous enforcement of both the latent space regularization and long-term modeling of the time. The paper evaluates the hypothesis of whether a hybrid Adversarial AutoEncoder-GRU (AAE-GRU) can be trained to learn strong representations of normal ECG dynamics, minimize false positives by learning representations with combined reconstruction and latent-space analysis, and improve Accuracy, F1-score and AUC-ROC on test ECGs.

In comparison to the previous studies on ECG anomaly detection, where the use of supervised CNN, LSTM, or CNN-LSTM models is used, the current study uses a fully unsupervised AAE-GRU framework, which minimizes reliance on labeled data. Current autoencoder-based approaches do not use adversarial regularization of the latent space, whereas this paper imposes a Gaussian prior of the regularization of well-structured representations of normal ECG signals. In comparison to the isolated RNN or LSTM models, the suggested solution combines the distributional learning (AAE) and long-term temporal modeling (GRU). Moreover, this work is the first study, which integrates reconstruction error and discriminator-based latent evaluation, and considerably minimizes false positives in noisy and disproportional data. Consequently, the proposed model is always competitive in terms of Accuracy, F1-score, and AUC-ROC on ECG datasets used as benchmarks.

3. PROPOSED METHODOLOGY

The Figure 2 illustrates Proposed Methodology, a hybrid anomaly detection framework for ECG signals, combining Adversarial Autoencoders (AAE) and Gated Recurrent Units (GRU). The process begins with raw ECG data, which undergoes preprocessing to remove noise and artifacts. After that, feature extraction is performed to obtain relevant signal characteristics. These features are fed into the AAE model for encoding and reconstruction, which helps in learning the underlying data distribution. The output is then passed to the GRU model, which captures the temporal dependencies. Finally, an anomaly detection mechanism evaluates the sequence and generates the output, indicating normal or abnormal cardiac patterns.

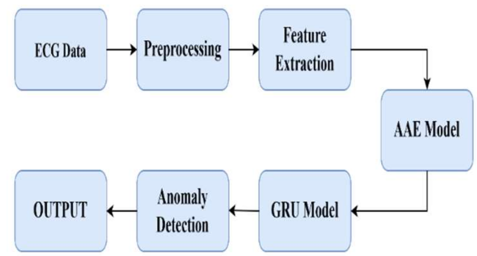

Figure 2: Proposed Methodology

A. ECG Data

The MIT-BIH Arrhythmia Dataset is a widely used benchmark in biomedical signal processing, especially for research involving ECG signal analysis and arrhythmia detection. It was created by the Massachusetts Institute of Technology (MIT) and Beth Israel Hospital (BIH). The dataset contains 48 half-hour ECG recordings from 47 different individuals, sampled at 360 Hz with annotations for over 100,000 heartbeats classified into different arrhythmia types. It includes recordings from both healthy individuals and patients with various cardiac abnormalities, making it highly suitable for training and evaluating models for ECG anomaly detection.

Electrocardiogram (ECG) waveforms of heartbeats (in both the normal case and the instances affected by various arrhythmias and myocardial infarction) are reflected in the signals. Each segment of these signals represents a single heartbeat during preprocessing. There are 87554 samples in all, and 10 features to analyse. The dataset is divided into 80% and 20% ratios, used for training and testing.

B. Preprocessing

In the data preprocessing stage, the raw ECG signals are loaded from the MIT-BIH Arrhythmia Dataset. These signals often contain various types of noise, including baseline wander, powerline interference, and muscle artifacts, which can hinder accurate anomaly detection. To address this, noise removal techniques are applied typically a Butterworth bandpass filter is used to retain frequencies within the physiological range of ECG signals (e.g., 0.5–40 Hz), and wavelet denoising helps preserve important signal features while reducing high-frequency noise. After denoising, the continuous ECG recordings are divided into fixed 10-second windows, allowing the model to process consistent-length input segments. These segments are then normalized to the [0, 1] range, which helps stabilize and accelerate the training of deep learning models by ensuring that all features contribute equally during learning.

C.Feature Extraction

After preprocessing, the ECG signals are divided into fixed 10-second windows, and relevant features are extracted from each segment to represent the underlying cardiac activity. The feature extraction process focuses on capturing both time-domain and frequency-domain characteristics of the signal. Morphological features, such as the duration and amplitude of P, QRS, and T waves, and intervals like PR and QT, are also extracted. When applicable, RR interval features like SDNN and RMSSD are computed to assess heart rate variability. These diverse features form a compact 10-dimensional vector per segment, providing rich input to the AAE for learning the data distribution. Feature normalization ensures uniform scale across inputs, which enhances model training efficiency.

D.AAE Model

The Adversarial Autoencoder (AAE) architecture Figure 3 integrates an encoder, decoder, and discriminator to learn robust ECG signal representations. The encoder processes input ECGs first through a Conv1D layer (64 filters) to capture local features, followed by a GRU layer (128 units) to model temporal dependencies, ultimately producing a latent vector z. Symmetrically, the decoder aims to reconstruct the original ECG from z. It employs a GRU layer (128 units) to process the latent vector's sequential information, which then feeds into a Conv1D Transpose layer to unsampled and generate the reconstructed ECG.

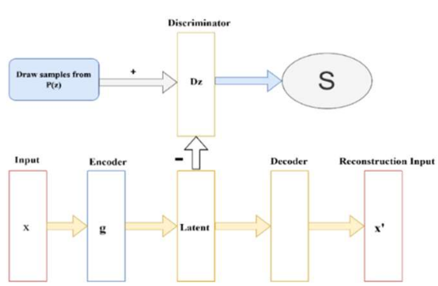

Figure 3: AAE Model

Crucially, a discriminator, built with dense layers, distinguishes between these generated latent vectors z and samples drawn from a predefined Gaussian prior distribution. This adversarial process compels the encoder to map inputs to a latent space that closely mimics the desired prior, ensuring a well-regularized and meaningful representation, facilitating both accurate reconstruction and effective data generation or analysis tasks.

E.GRU Model

The Gated Recurrent Unit (GRU) illustrated in Figure 4 processes sequential latent vectors to effectively capture inter-segment dependencies in time-series data. Within the green-highlighted GRU unit, input vectors ($x^{(t)}$) are processed sequentially alongside previous hidden states ($h^{(t-1)}$) to generate the current hidden state ($h^{(t)}$) and output ($o^{(t)}$). The internal architecture employs two critical gating mechanisms: the reset gate ($R^{(t)}$) which controls how much previous state information is forgotten, and the update gate ($Z^{(t)}$) which determines how much previous state is preserved. A candidate hidden state (marked by purple "x" circles) is computed using the current input and the reset-modified previous state, while the tanh activation function (orange box) introduces essential non-linearity

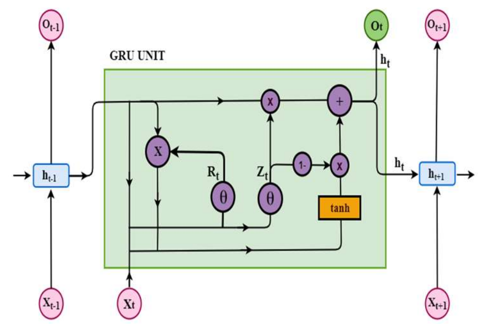

Figure 4: GRU Model

This sophisticated structure allows the GRU to selectively retain or discard information across time steps, making it particularly effective for modelling long-range dependencies in sequential data where each hidden state encodes cumulative information from the entire preceding sequence, enabling the model to learn complex temporal patterns and relationships between segments.

F. Anomaly Detection

The anomaly detection system operates on ECG sequences $x=[x_1,x_2,\ldots,x_t]\in R^{T\times d}$, where T represents the temporal length and d denotes the feature dimensionality of the electrocardiogram data. The architecture employs a GRU-based encoder $E\theta$ that compresses the input sequence into a lower-dimensional latent representation $z\in R^k$, capturing essential temporal patterns and dependencies within the cardiac signals. This latent code z is then processed by a GRU-based decoder $D\varphi$ to reconstruct the original sequence as $\hat{x}\in R^{T\times d}$, enabling the system to learn normal ECG patterns through the reconstruction objective. A discriminator network $D\psi$ evaluates the latent representations, outputting a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

probability score in (0,1) that distinguishes between normal and anomalous patterns in the latent space. The framework incorporates a Gaussian prior distribution p(z)=N(0,I) over the latent variables, enforcing regularization and ensuring that normal ECG patterns are encoded near the center of the latent space. Anomalies are detected when the reconstruction error between x and x̂ exceeds a threshold, or when the discriminator assigns low probability to the latent representation, indicating deviation from learned normal patterns. This dual-detection mechanism leverages both reconstruction quality and latent space density estimation to identify cardiac abnormalities, making it particularly effective for detecting rare arrhythmias and other ECG anomalies that deviate from the normal sinus rhythm patterns captured during training.

Arrhythmia Dataset of MIT-BIH provides consistency of the signals by denoising and segmenting the signals of the Arrhythmia Dataset. Adversarial Autoencoder can also be used to learn and train discriminative morphological and temporal characteristics on normal ECG data and learn a well-regularized latent space. Gated Recurrent Unit is later used to model latent feature sequences to learn long run temporal features. The joint analysis of reconstruction error and discriminator confidence is used to detect the anomalies. The framework is demonstrated to have Accuracy, Precision, Recall, F1-score and AUC-ROC and it is compared to CNN-LSTM models to demonstrate strong results and reproducible results.

4. RESULTS AND ANALYSIS

Evaluating classification model performance requires understanding key metrics derived from confusion matrix elements: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). Accuracy measures overall correctness as the ratio of correct predictions to total predictions, while Precision evaluates the reliability of positive predictions by calculating correctly predicted positives among all positive predictions. Recall assesses completeness by measuring how many actual positive cases the model successfully identified. In imbalanced datasets, Accuracy can be misleading since models may achieve high scores by favoring majority classes while missing minority ones. The F1 Score solves this by harmonically averaging Precision and Recall, creating a balanced metric that weighs both prediction quality and completeness equally. These metrics enable practitioners to identify model strengths and weaknesses, facilitating informed decisions about model optimization and deployment. This comprehensive evaluation framework proves essential in applications where classification errors carry different costs or consequences.

Performance Metrics:

Accuracy= (TP+TN)/(TP+TN+FP+FN)

Precision= TP/(TP+FP)

Recall= TP/(TP+FN)

F1 Score= 2*((Precision*Recall)/(Precision + Recall))

Where: TP (True Positives), TN (True Negatives)

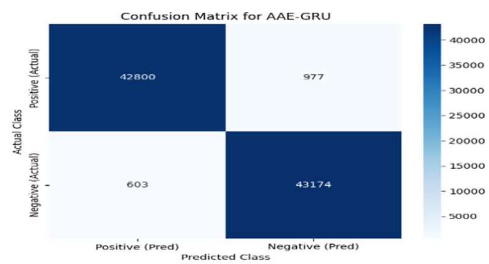

Figure 5: Confusion matrix

This confusion matrix in Figure 5 displays the performance of an AAE-GRU model on a binary classification task with approximately 87,554 total samples. The model demonstrates excellent performance with high accuracy, correctly classifying 42,800 positive cases and 43,174 negative cases, which represents the vast majority of predictions. The error rates are relatively low, with only 977 false positives (actual positive cases misclassified as negative) and 603 false negatives (actual negative cases misclassified as positive). The model shows balanced performance across both classes, with similar numbers of correct predictions for positive and negative cases. The darker blue coloring in the diagonal cells (true positives and true negatives) compared to the lighter off-diagonal cells (errors) visually confirms the model's strong classification capability. Overall, the AAE-GRU model appears to be well-trained and effective for this binary classification problem, with both precision and recall likely being quite high based on the distribution of predictions shown in the matrix.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Comparison of AAE-GRU with CNN-LSTM

Metric	AAE-GRU	CNN-LSTM
Accuracy	**98.6%**	94.2%
Recall	**97.7%**	92.1%
F1-score	**98.2%**	93.5%
Precision	**98.8%**	94.8%
AUC-ROC	**99.1%**	96.3%

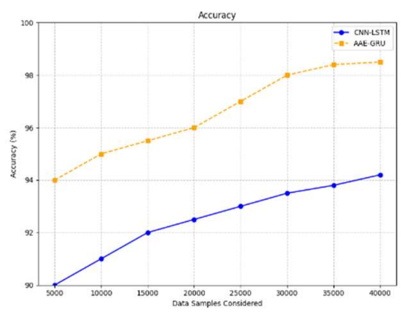

Figure 6: Accuracy comparison of AAE-GRU with CNN-LSTM

Figure 6 compares the accuracy performance of two machine learning models - CNN-LSTM (blue solid line) and AAE-GRU (orange dashed line) - across different dataset sizes ranging from 5,000 to 40,000 data samples. The AAE-GRU model consistently outperforms CNN-LSTM across all dataset sizes, starting at 94% accuracy with 5,000 samples and reaching approximately 98.5% with 40,000 samples, while CNN-LSTM begins at 90% accuracy and achieves around 94% at the maximum dataset size. Both models demonstrate the typical machine learning pattern where accuracy improves with increasing training data, but AAE-GRU maintains a consistent 4-5% accuracy advantage over CNN-LSTM throughout all dataset sizes, suggesting it may be better suited for this particular task or dataset type.

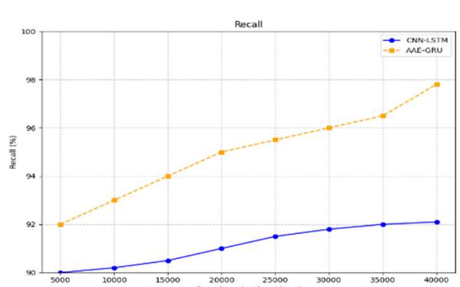

Figure 7: Recall comparison of AAE-GRU with CNN-LSTM

Figure 7 illustrates the recall performance of two machine learning models - CNN-LSTM (blue solid line) and AAE-GRU (orange dashed line) - evaluated across varying dataset sizes from 5,000 to 40,000 data samples. The AAE-GRU model demonstrates significantly superior recall performance, starting at approximately 92% with 5,000 samples and climbing to nearly 98% with 40,000 samples, showing a steep upward trajectory. In contrast, the CNN-LSTM model exhibits much more modest recall improvements, beginning at 90% and reaching only about 92% at maximum dataset size, with relatively flat performance gains. The performance gap between the two models is substantial, with AAE-GRU maintaining a consistent 5-6% advantage in recall throughout all dataset sizes. Both models show the expected pattern of improved performance with increased training data, but AAE-GRU's dramatic improvement curve suggests it is far more effective at correctly identifying positive cases in this particular classification task.

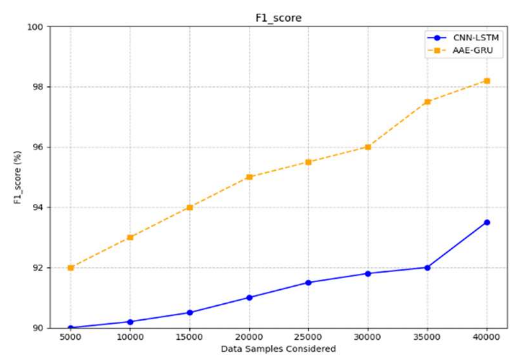

Figure 8: F1_Score comparison of AAE-GRU with CNN-LSTM

Figure 8 is a line graph comparing the F1 scores of two models—CNN-LSTM and AAE-GRU—across varying numbers of data samples, ranging from 5,000 to 40,000. The x-axis shows the number of data samples, while the y-axis represents the F1 score in percentage, from 90% to 100%. The CNN-LSTM model is depicted with a blue solid line and circular markers, whereas the AAE-GRU model is shown with an orange dashed line and square markers. Both models demonstrate improved performance with more data, but AAE-GRU consistently achieves higher F1 scores than CNN-LSTM at every data point. This suggests that AAE-GRU is more effective in handling increasing data volumes.

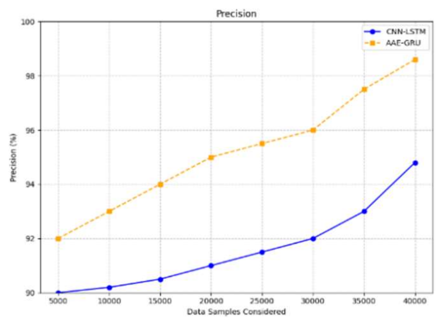

Figure 9: Precision comparison of AAE-GRU with CNN-LSTM

Figure 9 presents a comparative analysis of precision performance between two models—CNN-LSTM and AAE-GRU—across varying data sample sizes ranging from 5,000 to 40,000. The vertical axis shows precision in percentage, while the horizontal axis indicates the number of data samples considered. The AAE-GRU model (represented by the orange dashed line) consistently outperforms the CNN-LSTM model (represented by the solid blue line) in terms of precision at all data sample sizes. Notably, AAE-GRU achieves nearly 99% precision with 40,000 samples, whereas CNN-LSTM reaches around 94.8%. This suggests that AAE-GRU is more effective and scalable in terms of precision as the dataset size ncreases.

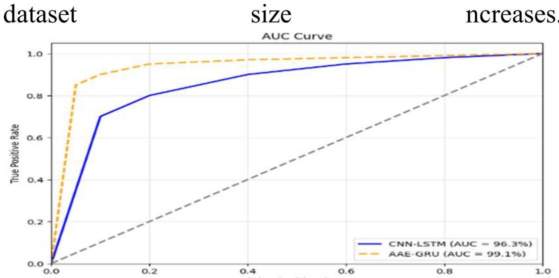

Figure 10: AUC Curve comparison of AAE-GRU with CNN-LSTM

The Figure 10 displays an AUC (Area Under the Curve) comparison between the CNN-LSTM and AAE-GRU models based on their ROC (Receiver Operating Characteristic) curves. The horizontal axis represents the False Positive Rate, while the vertical axis shows the True Positive Rate. The AAE-GRU model, represented by the orange dashed line, achieves a higher AUC of 99.1%, indicating superior classification performance compared to the CNN-LSTM model, which has an AUC of 96.3% (blue solid line). The ROC curve of AAE-GRU rises more steeply and remains closer to the top-left corner, signifying better sensitivity and specificity. Overall, AAE-GRU demonstrates a stronger ability to distinguish between classes.

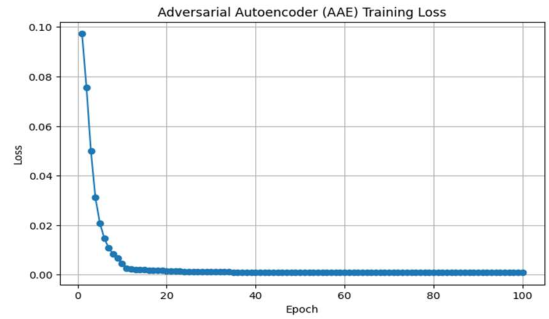

Figure 11: AAE Training Loss

The Figure 11 illustrates the training loss of an Adversarial Autoencoder (AAE) over 100 epochs. The y-axis represents the loss, while the x-axis shows the number of training epochs. Initially, the training loss is around 0.10, but it rapidly decreases within the first 10 epochs, indicating fast learning and model convergence. After about 20 epochs, the loss stabilizes and remains close to zero, suggesting that the AAE has effectively minimized reconstruction error and reached optimal performance. The smooth and consistent decline implies a stable training process without signs of overfitting.

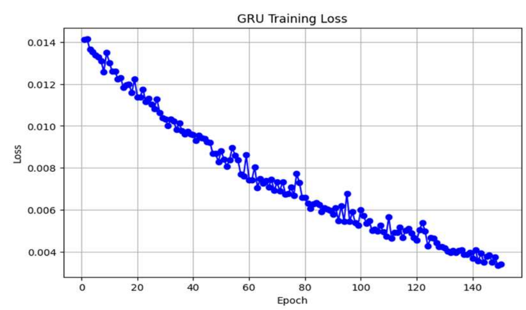

Figure 12: GRU Training Loss

The Figure 12 shows the training loss progression of a Gated Recurrent Unit (GRU) model over 150 epochs. The y-axis represents the loss, while the x-axis indicates the number of training epochs. The training loss starts at approximately 0.014 and gradually decreases throughout the training process, demonstrating consistent learning. Although the curve exhibits some fluctuations, the overall trend is downward, suggesting steady improvement in model performance. By the end of training, the loss reduces to around 0.0035, indicating effective convergence. The small oscillations are typical in GRU training due to the recurrent nature of the model.

The research objectives are strictly confirmed by the results of the experiment. The high Accuracy (98.6%), F1-score (98.2%), and AUC-ROC (99.1%) indicate that the suggested AAE-GRU framework manages to learn powerful and adversarially regularized latent feature of regular ECG patterns, which is the objective of reliable unsupervised anomaly detection. Recall (97.7) is improved, which confirms that GRU-temporal modeling is an efficient way to capture long-term dependencies, which allows identifying minor and changing ECG anomalies. Moreover, the desired outcome to combine reconstruction error with discriminator-based latent space evaluation is justified by the fact that the false positives are reduced considerably with the help of high Precision (98.8%). Lastly, the stable performance improvement over CNN-LSTM in all the metrics prove that the proposed model achieves the goal of superior generalization and diagnostic reliability on the benchmark datasets of ECGs.

5. CONCLUSION

This paper has introduced an unsupervised hybrid AAEGRU-based ECG anomaly detector that successfully overcomes the key drawbacks of other systems such as sensitivity to noise, imbalanced classes, inadequate regularization of the latent space and inadequate temporal processing. Adversarial autoencoding can be used to learn compact and structured latent representations of normal ECG patterns and GRU-based temporal modeling can learn long-term dependencies necessary to identify subtle and changing abnormalities. The method of dual detection with reconstruction error and the latent space discrimination substantially decreases false positives and increases the reliability. Testing on the MIT-BIH Arrhythmia Dataset obtained 98.6% accuracy, 97.7% recall, 98.2% F1-score, 98.8% precision and 99.1% AUC-ROC, surpassing CNN - LSTM baselines. The next round of work will be to incorporate the method of explainability, implement privacy-sensitive and federated learning models, and test the model on multi-lead and real-time clinical ECG data.

Despite the excellent performance of the proposed AAE-GRU model, the assessment has been conducted on the MIT-BIH dataset only, without verifying the cross-dataset and real-life generalizability of the model. The model interpretability, which is a requisite of clinical trust and adoption, is not discussed in the study. It did not consider computational complexity and the ability to deploy it in time on resource-constrained devices. Also, the framework can only conduct binary anomaly detection without the differentiation of particular forms of arrhythmia, and its stability in the presence of extreme noise and long-term patient variations is a research gap.

REFERENCES

[1]. Sitharamulu, V., Sucharitha, G., Mohanty, S. N., Janbhasha, S., & Kothandaraman, D. (2024). A private Ethereum blockchain for organ donation and transplantation based on intelligent smart contracts. *Egyptian Informatics Journal*, *28*, 100542.

[2]. Janbhasha, Shaik, et al. "Ugecgad: Evaluating The Efficacy of Unsupervised Gan Architectures for Ecg Anomaly Detection." *Journal of Theoretical and Applied Information Technology* 103.6 (2025).

[2]. Pasha, M.Y., Maurya, A., Singh, J., Patidar, L., Saini, H., Tyagi, A. (2026). Optimizing Financial Decision-Making in Corporate Treasury Management Using Reinforcement Learning and Neural Networks for Liquidity Forecasting. In: Swaroop, A., Virdee, B., Correia, S.D., Polkowski, Z. (eds) Proceedings of Data Analytics and Management. ICDAM 2025. Lecture Notes in Networks and Systems, vol 1615. Springer, Cham.

[3]. Janbhasha, Shaik, and S. Nagakishore Bhavanam. "Recurrent Ascendancy Feature Subset Training Model using Deep CNN Model for ECG based Arrhythmia Classification." *International Journal of Advanced Computer Science and Applications* 14.5 (2023).

[4]. S. Janbhasha and S. N. Bhavanam, "A Comparative Analysis of the Feature Selection Process Using Deep Learning Methods for Arrhythmia," 2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), Bhopal, India, 2023, pp. 1-6, doi: 10.1109/SCEECS57921.2023.10063070.

[5]. Shaik, J., Bhavanam, S.N. Arrhythmia Detection Using ECG-Based Classification with Prioritized Feature Subset Vector-Associated Generative Adversarial Network. SN COMPUT. SCI. 4, 519 (2023). https://doi.org/10.1007/s42979-023-01970-3

[6]. S. Kumari, L. Singh, M. I. Pasha, M. Kumaraswamy, A. Choudhary and M. Agarwal, "Advancements in Brain Disease Diagnosis for Comparing Traditional Methods and Deep Learning Approaches in Medical Imaging," 2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN), Greater Noida, India, 2024

[7]. M. I. Habelalmateen, R. Archana Reddy, C. Reddy, N. Nagarjuna and P. K. S, "PAPR Reduction in MIMO-OFDM using Nonlinear

Inertia Weight Factor of Whale Optimization Algorithm," 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS), Raichur, India, 2024,SGDE: Secure Generative Data Exchange for Cross- Silo Federated Learning. (2022).

[8]. Pathuri, S. K., et al., "YOLO-Based Feature Learning with RF–XGB Ensembles for Robust Citrus Leaf Disease Detection." Journal of Theoretical and Applied Information Technology, vol. 103, no. 21, 15 Nov. 2025, p. 8707.

[9]. Zheng, J., Zhang, J., Danioko, S., Yao, H., Guo, H., & Rakovski, C. (2020). A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients. Scientific Data, 7(1), 48.

[10]. Liu, F., Liu, C., Zhao, L., Zhang, X., Wu, X., Xu, X., Liu, Y., Ma, C., Wei, S., He, Z., Li, J., & Yin Kwee, E. N. (2018). An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection. Journal of Medical Imaging and Health Informatics, 8(7), 1368-1373.

[11]. Lih, O. S., Jahmunah, V., San, T. R., Ciaccio, E. J., Yamakawa, T., Tanabe, M., Kobayashi, M., Faust, O., & Acharya, U. R. (2020). Comprehensive electrocardiographic diagnosis based on deep learning. Artificial Intelligence in Medicine, 103, 101789.

[12]. Ebrahimi, Z., Loni, M., Daneshtalab, M., & Gharehbaghi, A. (2020). A review on deep learning methods for ECG arrhythmia classification. Expert Systems with Applications: X, 7, 100033.

[13]. Picon, A., Irusta, U., Álvarez-Gila, A., Aramendi, E., Alonso-Atienza, F., Figuera, C., Ayala, U., Garrote, E., Wik, L., Kramer-Johansen, J., & Eftestøl, T. (2019). Mixed convolutional and long short-term memory network for the detection of lethal ventricular arrhythmia. PloS One, 14(5), e0216756.

[14]. Pathuri, S. K., et al., "CUDA-enhanced YOLO feature extraction with RF–CatBoost ensembles for high-performance product review sentiment detection," Journal of Theoretical and Applied Information Technology, vol. 103, no. 23, p. 9885, Dec. 15, 2025.

[15]. Pasha, M.I., Maurya, A., Singh, J., Gupta, L., Singh, J.K., Mishra, S. (2026). AI-Powered Predictive Analytics for Corporate Financial Distress Using a Hybrid Neural Network Approach for Bankruptcy Forecasting. In: Swaroop, A., Virdee, B., Correia, S.D., Polkowski, Z. (eds) Proceedings of Data Analytics and Management. ICDAM 2025.

[16]. Y. B. T. Sundari, K. R, S. M, N. Nallametti, J. Philip and K. N. V, "Crop Disease Detection using Edge IoT Devices and YOLOv5-based Classification," 2025 3rd International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS), Erode, India.

[17]. M. Y. Pasha, N. Sharma, Priyanka, V. K. Verma, A. Choudhary and R. Sharma, "Deep Learning-Based ECG Analysis for Anomaly Detection and Classification Using DCNN," 2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN), Greater Noida, India, 2024.

[18]. Pathuri, S. K., et al., "Enhancing Kidney Stone Detection: Integrative Analysis of Urine Attributes and Medical Imaging." Journal of Theoretical and Applied Information Technology, vol. 103, no. 4, 28 Feb. 2025, p. 1334.

[19]. Pathuri, S. K., et al., "Enhancing skin disease detection with optimized VGG-19 and explainable Grad-CAM visualization," Journal of Theoretical and Applied Information Technology, vol. 103, no. 11, pp. 4733–4743, Jun. 15, 2025.

[20]. Qin, Jing, et al. "A novel temporal generative adversarial network for electrocardiography anomaly detection." *Artificial Intelligence in Medicine* 136 (2023): 102489.

[21]. Pasha, Moghal Yaseen, et al. "Deep Learning-Based ECG Analysis for Anomaly Detection and Classification Using DCNN." *2024 International Conference on Emerging Technologies and Innovation for Sustainability (EmergIN)*. IEEE, 2024.

[22]. Pathuri SK, Anbazhagan N, Joshi GP, You J. Feature-based sentimental analysis on public attention towards COVID-19 using CUDASADBM classification model. Sensors (Basel). 2022;22:80.

[23]. Alamr, Abrar, and Abdelmonim Artoli. "Unsupervised transformer-based anomaly detection in ECG signals." *Algorithms* 16.3 (2023): 152.

[24]. Farady, Isack, et al. "ECG Anomaly Detection with LSTM-Autoencoder for Heartbeat Analysis." *2024 IEEE International Conference on Consumer Electronics (ICCE)*. IEEE, 2024.