IRJSE International Research Journal of Science & Engineering

Int. Res. J. of Science & Engineering, Special Issue A14, June, 2024
ISSN: 2322-0015

SJIF Impact Factor 6.70
https://doi.org/10.5281/zenodo.14198945

RESEARCH ARTICLE OPEN ACCESS

Anomaly Detection in ECG Signals: A Comparative Study of Autoencoder, GAN and Hybrid GAN-AE Models

Adithya S[1] and Prathibha Prakash[2]

[1]Department of Mathematics, Amrita Vishwa Vidyapeetham,
[2]Department of Computer Science and Engineering, Amrita Vishwa Vidyapeetham,
Email: [1]adithya7122@gmail.com , [2]prathibhaprakash@am.amrita.edu

Manuscript Details

Available online on https://www.irjse.in
ISSN: 2322-0015

Editor: Dr. Arvind Chavhan

Cite this article as:
Adithya S and Prathibha Prakash. Anomaly Detection in ECG Signals: A Comparative Study of Autoencoder, GAN and Hybrid GAN-AE Models, *Int. Res. Journal of Science & Engineering,* 2024, Special Issue A14: 165-172. https://doi.org/10.5281/zenodo.14198945

Article published in Special issue of National Conference on Machine Learning and Data Science (NCMLDS-2024)" organized by College of Computer Science and Information Technology (COCSIT) Ambajogai Road, Latur, Maharashtra, India on date April 16th to 17th 2024

Abstract

The electrocardiogram (ECG) is a regular test that measures the activity of your heart. The ECG measures heart activity using electrical signals. The ECG can indicate a variety of cardiac diseases, including arrhythmia (an irregular heartbeat). Arrhythmia diagnosis is based on distinguishing normal and irregular heartbeats and appropriately classifying them with ECG morphology. Anomaly detection in electrocardiogram (ECG) signals is crucial for early cardiovascular disease diagnosis. This study compares autoencoders (AEs), generative adversarial networks (GANs), and a novel hybrid GAN-AE model for detecting ECG abnormalities. The hybrid model, which was trained exclusively on normal ECG signals, is excellent at recognizing subtle patterns. The evaluation computes reconstruction loss for test ECG data and applies a human threshold to detect problems. The hybrid GAN-AE outperforms, with higher accuracy. The hybrid model, which combines Gans' creative power with AEs' reconstruction skills, can recognize sensitive ECG patterns. The study emphasizes the need of early anomaly detection in ECG signals for proactive healthcare.

Keywords: Anomaly Detection, ECG signals, Autoencoder, GAN, Hybrid GAN-AE.

1. Introduction

Electrocardiogram (ECG) signals are time series data sequences that provide a graphical representation of electrical impulses from the heart [1]. The ECG measures cardiac activity in the form of electrical signals from the heartbeat. During testing, data is collected from a number of electrodes placed on the chest skin. The great majority of clinicians prefer to use ECG signals as their primary screening tool for detecting and diagnosing cardiac issues. A regular ECG report reveals a consistent heart rhythm, but an abnormal ECG has no consistent rhythmic pattern. An abnormal ECG indicates heart muscle injury or rhythmic difficulties in the heart. Traditionally, medical practitioners decide if an ECG report is normal or abnormal by observing the rhythmic patterns of the ECG signal and applying their expertise and experience. However, the privacy restrictions controlling personal healthcare records, the need for skilled expert annotators, the cost of collecting and annotating ECGs, and the unpredictable nature of ECGs make dataset development a tough task. Furthermore, trained healthcare experts or doctors may not always be available in emergencies, particularly in remote areas. In this context, automatic arrhythmia prediction using ECG signals is a critical component of building artificial intelligence applications to treat cardiac sickness. A large annotated ECG dataset is also necessary for the creation of a reliable and scalable machine learning or deep learning-based model for detecting ECG anomalies. As a result, designing and implementing appropriate algorithms for automatically detecting arrhythmia from ECGs is becoming increasingly important.

In recent years, researchers have focused on abnormality predictions from ECG data using machine learning and deep learning models. DL is a subset of machine learning, and most DL methods make numerous promises when compared to traditional ML, including higher accuracy, greater flexibility, stronger generalization, and less reliance because DL methods learn the inherent features of the data automatically, eliminating the need for feature engineering to extract and select appropriate features for the given task. As a result, deep learning (DL) methods surpass traditional machine learning (ML) methods in terms of both learning ability and accuracy. The most difficult challenge is dealing with real-world time series data [2], which often contains partial or missing values, as well as noise. DL techniques have been effectively used to a number of time series data processing applications, including electrocardiogram (ECG), electroencephalogram (EEC), and time sequence dependent phenomena.

This research proposes a hybrid GAN-AUTOENCODER architecture for detecting arrhythmias and anomalous ECG signals.

Automated ECG abnormality detection has various challenges, including:

- The model has a limited number of anomalies to train on.
- Imbalance in data, with normal ECGs outnumbering anomalies.

The proposed method addresses the aforementioned issues by training the model with only normal ECG signals (no anomalous ECG signals are required during the training phase) and attaining high accuracy during the testing phase using both normal and anomalous test ECG signals. Thus, data imbalances and a scarcity of annotated samples are handled.

2. Related Work

Automatic anomaly identification using time series data such as electrocardiograms (ECGs) and electroencephalograms (EEGs) via deep learning (DL) algorithms has risen in popularity in recent years. Several researchers have proposed supervised [3], unsupervised, and semi supervised deep learning methods to identify irregular heartbeats.

The study [4] provides a complete review and evaluation of the usage of deep learning techniques in the context of heartbeat detection using electrocardiogram (ECG) signals. Accurate heartbeat detection is crucial for a wide range of therapeutic applications, and recent improvements in deep learning have shown promising results in boosting the process's precision. Addresses challenges that arise when using deep learning to heartbeat detection, such as the need for large annotated datasets and deep model interpretability.

Imbalanced time series are widespread in industrial applications where the number of normal samples far exceeds the number of aberrant cases. Traditional machine learning algorithms, such as support vector machines and convolutional neural networks, struggle to attain high classification accuracies for class-imbalanced problems because they are concerned with the correctness of the majority class. To solve this issue, this study [5] proposes a novel anomaly detection approach based on generative adversarial networks (GAN). Limitations such as susceptibility to imbalance ratios, computing complexity, and interpretability issues are mentioned.

In paper [6] Yahoo! Webscope S5 dataset was used to test machine learning algorithms for detecting anomalies in time-series data. They used GANs to create synthetic data for detecting anomalies in time series data. Despite the promising results, the proposed method has significant limitations, such as data dependency, computational complexity, and threshold sensitivity.

The approach suggested in this paper does not require aberrant signals during the autoencoder training phase; training is done with only normal ECG signals, which is a significant advantage. In real-world scenarios, annotated abnormal signals may not always be available during the training process when constructing an ECG prediction model, but it is required to predict both normal and abnormal signals. In this instance, the solution we propose will work.

3. Dataset

The original ”ECG5000” dataset is a 20-hour ECG acquired from Physionet. It is record ”chf07” in the BIDMC Congestive Heart Failure Database (chfdb). ”Goldberger AL, Amaral LAN, Glass L, Hausdorff JM, Ivanov PCh, Mark RG, Mietus JE, Moody GB, Peng C-K, Stanley HE” was the initial publication.PhysioBank, PhysioToolkit, and PhysioNet: Com-ponents of a New Research Resource for Complex Physiologic Signals.” ”101(23) Circulation.” The data was pre-processed in two steps:

- extract each heartbeat.
- interpolate each heartbeat to equal length.

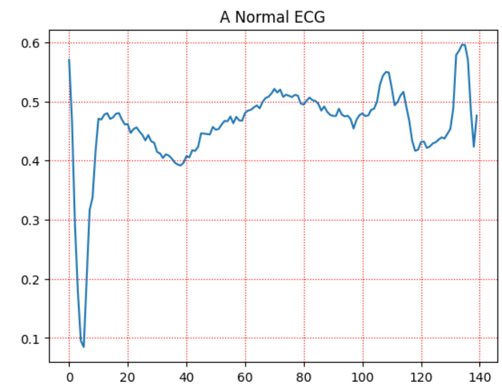

Figure 1. Normalous ECG

This dataset was initially used in the publication” A general framework for never-ending learning from time series streams”, DAMI 29(6). Then 5,000 heartbeats were chosen at random. The patient suffers from severe congestive heart failure, and the class values were derived using automated annotation. This data collection includes 5000 ECGs with 140 data points each.

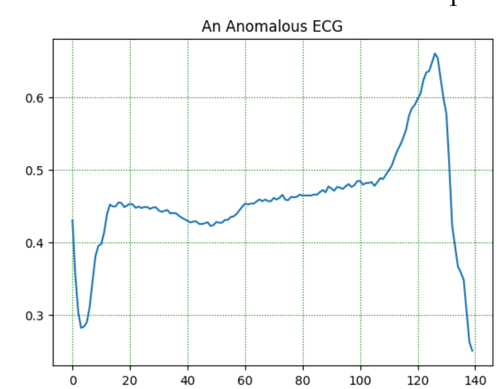

Figure 2. Anomalous ECG

4. Proposed Model

This section defines Autoencoder, GAN, and Hybrid GAN-AE briefly

4.1. Autoencoder

The autoencoder (AE) is a neural network design that can learn the structure of a dataset and generate a compressed representation of the input data. There are several types of autoencoder architectures available for learning and ensuring that the compressed representation faithfully duplicates the original data input. This sort of learning does not aim to create a model capable of making accurate predictions on new data, but rather to explore data attributes and find features or trends. Autoencoder neural networks are used to learn how to represent incoming data. The autoencoder's principal goal is to reproduce the output as nearly as possible to the input data. It consists of two components: encoder and decoder.

The encoder transforms incoming data $x \in R^m$ to create a compressed representation known as the bottleneck layer or latent space. This representation $z \in R^n$ (where $n < m$) contains important features from the original input. As a result, the data compression approach is implemented in the encoder part.

The second component (the decoder) applies the encoder's inverse changes to the original input data to produce the decoded output $x' \in R^q$ (where $q > n$). The decoder portion reconstructs the input signal from the encoder's latent space representation (z)into a dimension that is as near to the original input data as feasible.

The dimensions of the input signal x, latent space representation z, and decoded output x' are denoted by m, n, and q, respectively. Please keep in mind that the size q of the decoded output for the present task equals the dimension m of the input data.

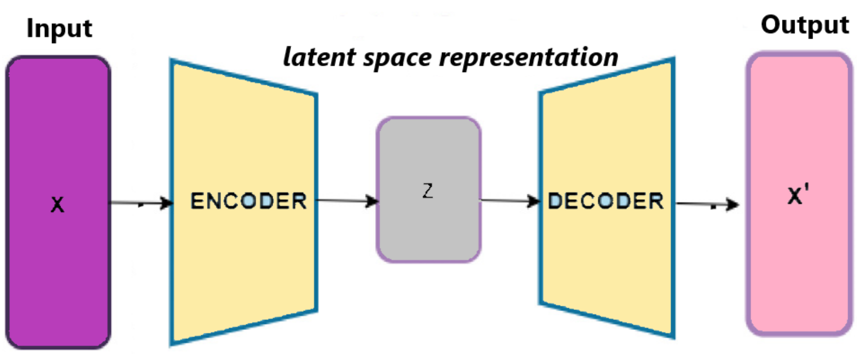

Figure 3. Architecture of Autoencoder

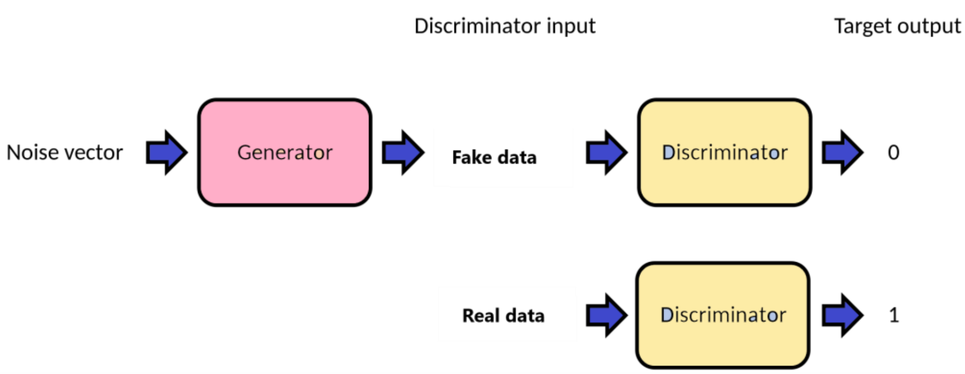

Figure 4. Architecture of GAN

4.2. Generative Adversarial Network (GAN)

A GAN, or Generative Adversarial Network, is a form of artificial intelligence algorithm used for unsupervised machine learning. It has evolved into a useful technique for detecting abnormalities in a wide range of fields, including image processing, network traffic analysis, and fraud detection. They are ideal for detecting anomalous patterns or outliers because of their ability to capture complex data distributions and identify deviations from the mean.

A GAN consists of two competing neural networks: a generator and a discriminator. The generator is tasked with making realistic samples that appear to be actual data, while the discriminator is tasked with discriminating between real and generated samples.

During training, the generator and discriminator engage in a minimax game in which the generator strives to enhance its ability to generate realistic samples while the discriminator constantly improves its discrimination abilities. This adversarial training strategy encourages the generator to understand the underlying patterns and distributions of normal data, helping it to capture the properties of normal data points more effectively.
The discriminator's output, which represents the possibility of a sample being actual, serves as an anomaly score. Low anomaly numbers are regarded normal, however high anomaly scores are classified as likely anomalies.

4.3. Hybrid GAN-AE

Hybrid GAN-AE for anomaly detection is an effective strategy that combines the strengths of generative adversarial networks (GANs) and autoencoders (AEs) to locate anomalies in a variety of data sets.

GANs excel at capturing the underlying data distribution and producing realistic samples, whereas AEs excel in data reconstruction and learning latent representations that capture the essential features of the data. When these two procedures are combined hybrid GAN-AE models can outperform conventional methods for anomaly detection. In a hybrid model, the autoencoder functions as a generator, producing realistic samples.

The architecture of a hybrid GAN-AE model typically consists of two main components:

- Autoencoder
- Discriminator

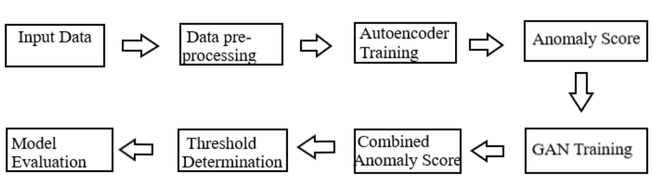

Figure 5. Block Diagram of Hybrid GAN-AE

4.4. Data Preprocessing

Normalizing the data ensures that all attributes have the same scale, which can improve the convergence and performance of machine learning models, particularly neural networks.

4.5. Data Splitting

Initially, the dataset was processed to separate features from labels. The label was taken from the dataset's last column, and the feature matrix was constructed from the remaining columns. The dataset was then divided into training and test sets. The split was conducted with a 20% test size.

4.6. Encoder

The stated encoder module consists of two dense layers. Each ECG data set included 140 continuous time steps of the heartbeat rhythm. This normal ECG time series data is passed to the first dense layer (140 × 1). To capture the data's temporal dependency. The first layer contains 32 hidden neurons. The sequences generated by these hidden states are then triggered using the Rectified Linear Unit (ReLU) method. These encoded representations include the non-linear temporal features learned from the ECG data so far. The sequential data is further encoded by a second dense layer with 16 hidden neurons. This module generates the final encoded latent space representation of the original ECG signal.

4.7. Decoder

The Decoder block, like the Encoder block, consists of two dense layers. It is actually the mirror image of the encoder module, with the decoder layers stacked in the opposite order as the encoder block. The purpose of this decoder module is to recover the original data from the encoded latent space representation signals generated by the dense based encoder module. Because the decoder module is also a feed forward network, its first layer takes the encoded representation from the encoder module's last layer. The initial dense layer consists of 16 hidden neurons. In this case, the activation function ReLU (Rectified Linear Unit) is used to extract nonlinear temporal information from the signal.

4.8. Discriminator

The discriminator is constructed as a sequential model with four layers. The first layer is a Flatten layer, which converts the input data with the shape (140, 1) into a one-dimensional array. Following flattening, two dense layers of 32 and 16 neurons each use the rectified linear unit (ReLU) activation function. The final dense layer is made up of one neuron with a sigmoid activation function and serves as the output layer for binary classification results (real or fake). The model is then built with the Adam optimizer, binary crossentropy loss, and accuracy as the evaluation metric. This discriminator is intended to check the legitimacy of input data in the GAN framework, distinguishing between actual and produced samples.

4.9. Combine Autoencoder and GAN

A combined model is created by integrating an autoencoder and a discriminator into a Generative Adversarial Network (GAN). It sets the autoencoder's trainable attribute to False, guaranteeing that it remains fixed throughout training. The GAN input's form corresponds to the latent dimensions. The autoencoder is then applied to the GAN input, and the results are sent to the discriminator. The final GAN model, which includes the autoencoder and discriminator, is built with the Adam optimizer, binary cross entropy loss, and accuracy as the evaluation measure. This integrated model allows the autoencoder to produce samples and the discriminator to distinguish between real and generated samples, both of which contribute to GAN training.

4.10. Autoencoder Training

The fitted approach is used to train the autoencoder model on normal training data. The test data is reconstructed using a trained autoencoder. The average squared error between the reconstructed and original test data is then determined. The resulting loss represents reconstruction errors.

4.11. Training GAN

The procedure involves generating normal samples with random noise, taking random batches of normal data, labeling the real and created samples, and training the discriminator to distinguish between them. The generator is also trained to provide data that can trick the discriminator. This adversarial training method is repeated for several epochs. The printed progress shows the discriminator loss, accuracy, and generator loss, enabling for monitoring and evaluation of the GAN's performance in creating realistic data and detecting abnormalities.

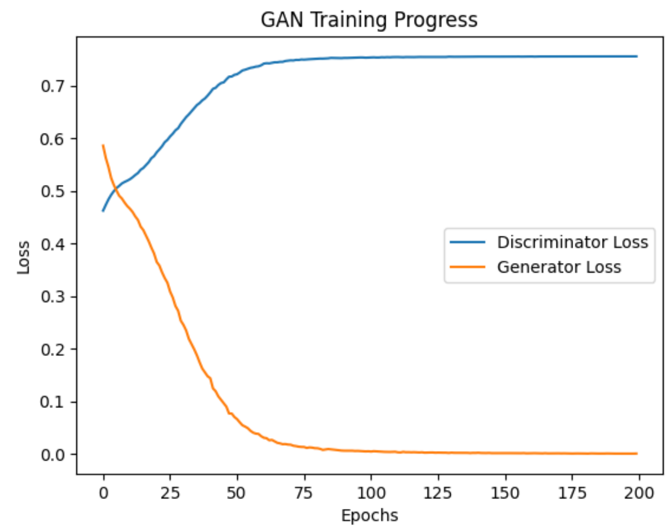

Figure 6. Discriminator loss vs Generator loss

4.12. Combined Anomaly Score

We create fake data for testing using random noise. The resulting data is then flattened to compute the reconstruction error (MSE) between the original and autoencoder-derived data. Furthermore, discriminator scores are collected from the GAN. An anomaly score is calculated by combining the reconstruction error and discriminator scores with a weight parameter (alpha) that the user defines. This combined score gives a comprehensive measure of anomaly detection, taking into account both autoencoder reconstruction and GAN discrimination.

4.13. Threshold selection

After training the model, the threshold value is determined and predictions are made based on it, so that reconstructed ECG signals with loss values less than the threshold value are classified as normal, and those with loss values greater than the threshold value are classified as anomaly ECGs. GAN discrimination.

4.14. Hyperparameter Setting

Table 1. Overview of training hyper-parameters

Batch size	512
Activation Function	Relu
No. of epoch	200
Optimizer	Adaptive momentum
Loss	L1 loss

5. Experimental Evaluation

To analyze and compare the performance of these three models, we need evaluation metrics. Here we choose two evaluation metrics which are accuracy and precision, recall and F1 score.

Accuracy: The amount of accuracy indicates how accurate a model is overall in its predictions. It determines how many samples were correctly predicted out of all the samples.

For binary classification tasks:

$$accuracy = \frac{TP+TN}{(TP+TN+FP+FN)}$$

where,

TP: True positives (correctly predicted positive cases)

TN: True negatives (correctly predicted negative cases)

FP: False positives (incorrectly predicted positive cases)

FN: False negatives (incorrectly predicted negative cases)

Precision: The percentage of optimistic forecasts that come true in reality. It indicates the frequency with which the model accurately predicts a positive case.

$$precision = \frac{TP}{(TP+FP)}$$

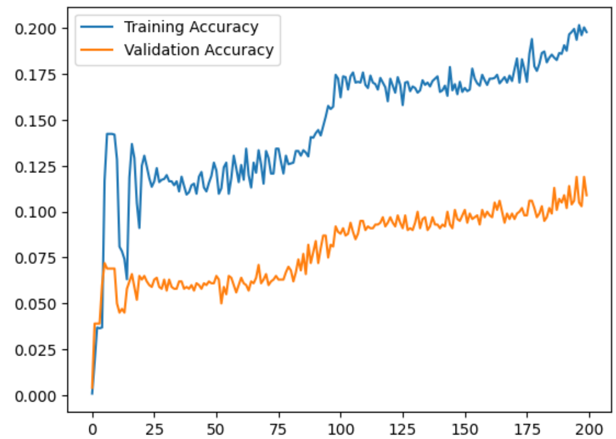

Figure 7. Training and Validation Accuracy of Autoencoder

Recall: Recall, also known as sensitivity or true positive rate, is a classification performance statistic. It quantifies a model's capacity to accurately identify all relevant cases, or the fraction of actual positive examples correctly identified by the model.

$$recall = \frac{TP}{(TP+FN)}$$

F1 score: The F1 score can be thought of as the weighted average of precision and recall, with a maximum value of 1 indicating that both precision and recall are perfect.

$$f1 = \frac{(2 \times precision \times recall)}{precision + recall}$$

5.1. Experimental Results

Table 2. Results

Model	Accuracy	Precision	Recall	F1-sore
Autoencoder	0.94	1.0	0.947	0.972
GAN	0.85	0.5596	0.998	0.717
Hybrid GAN-AE	0.99	1.0	0.99	0.99

The results reveal that the hybrid model outperformed the results of the individual models.

8. Conclusion

In this paper, we investigated three distinct models for anomaly detection on ECG5000 dataset, namely Autoencoder, GAN, and Hybrid GAN-AE. The major goal was to compare their performance in detecting abnormalities in complicated datasets. Our experimental results showed that the Hybrid GAN-AE model outperformed both the standalone Autoencoder and GAN models in terms of accuracy.

The Hybrid GAN-AE model demonstrated a novel capacity to grasp complex patterns within the data, resulting in higher anomaly detection accuracy by utilizing the capabilities of both generative and autoencoding architectures. This discovery highlights the value of integrating generative and autoencoding techniques to improve the overall durability and flexibility of the anomaly detection system.

Finally, our research demonstrates the Hybrid GAN-AE model as a powerful tool for anomaly detection.

9. References

1. James S Fleming, "Interpreting the electrocardiogram", Springer Science & Business Media, (2012).
2. John Cristian Borges Gamboa, "Deep learning for time-series analysis", arXiv preprint arXiv:1701.01887, (2017).
3. H Altay Guvenir, Burak Acar, Gulsen Demiroz, and Ayhan Cekin, "A supervised machine learning algorithm for arrhythmia analysis", In Computers in Cardiology 1997, pages 433–436. IEEE, (1997).
4. Fatma Murat, Ozal Yildirim, Muhammed Talo, Ulas Baran Baloglu, Yakup Demir, and U. Rajendra Acharya," Application of deep learning techniques for heartbeats detection using ecg signals-analysis and review", Computers in Biology and Medicine, 120:103726, (2020).
5. Rohit Raturi, Abhishek Kumar, Narayan Vyas, and Vishal Dutt,"A novel approach for anomaly detection in time-series data using generative adversarial networks", In 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), pages 1352–1357. IEEE, (2023).
6. Haoyuan Li and Yifan Li,"Anomaly detection methods based on gan: a survey", Applied Intelligence, vol. 53, no. 7, (2023), pp. 8209-8231.
7. A Anandharaj and P Bagavathi Sivakumar,"Anomaly detection in time series data using hierarchical temporal memory model", In 2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA), pages 1287–1292. IEEE, (2019).
8. GK Nischitha, S Manishankar, Phani Deshpande, and A Anoop," A cnn based anomaly detection system for real time fog based application", In 2021 Asian Conference on Innovation in Technology (ASIANCON), pages 1–7. IEEE, (2021).
9. Gayathiri RR and Nirmala Devi M," Unsupervised anomaly detection to handle imbalanced datasets using auto encoders for ecg signal classification", In 2023 Third International Conference on Secure Cyber Computing and Communication (ICSCCC), pages 263–268, (2023).
10. Maya Manish Kumar and GR Ramya," Performance comparison of anomaly detection algorithms", In Inventive Communication and Com-putational Technologies: Proceedings of ICICCT 2021, pages 761–776. Springer, (2022).
11. Sabari KK, Saurabh Shrivastava, and Sangeetha. V, "Anomaly-based intrusion detection using gan for industrial control systems", In 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pages 1–6, (2022).
12. K Devika, Dwarikanath Mahapatra, Ramanathan Subramanian, and Venkata Ramana Murthy Oruganti, "Outlier-based autism detection using longitudinal structural mri", IEEE Access, 10:27794–27808, (2022).

Conflicts of interest: The authors stated that no conflicts of interest.

Publisher's Note
IJLSCI remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Correspondence and requests for materials should be addressed to Patil SS.

Peer review information
IRJSE thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at https://www.irjse.in/reprints