SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Access

Anomaly Detection in Electrocardiogram Using Fully Convolutional Neural Network for Cardiac Disease Diagnosis

Pushpavalli K[1], Dr. Arasakumar M[2], Dr. S. Balaji[3]

[1]Research Scholar Department of Information Technology, Annamalai University. E-mail: pushpavalli.k@gmail.com
[2]Assistant Professor Department of Information Technology, Annamalai University. E-mail: dr.arasakumar.m@gmail.com
[3]Professor Department of Computer Science and Engineering, Panimalar Engineering College. E-mail: balajiit@gmail.com

Abstract
The overall aim of the research is to design a healthcare monitoring system for cardiac patients encompassing an efficient sensor network for collecting and transmitting physiological data captured from multiple health care devices to a cloud repository and a reliable deep learning based physiological data processing algorithms for detecting anomaly. Anomaly detection (aka outlier analysis) is a process to identify data points in a sequence of observations that deviate from a waveform's normal behavior. This paper presents a detailed review and experimental results with analysis of the anomaly detection process using fully convolutional neural networks. The potential of Encoder-Decoder architecture AutoEncoder (AE) and its variants in detecting anomalies in Echocardiogram (ECG) data is explored. The problem of detecting anomalies in the ECG data for cardiac disease analysis is solved using unsupervised learning technique. The process of AE-based anomaly detection involves compressing normal data using the AE into a latent space with a lower dimension than the original data, restoring the data, and calculating the difference between the restored and original data. The suitability of the Variational Autoencoder (VAE) and the challenges in fixing an optimal threshold value for categorizing anomaly samples were analyzed in this paper.

keywords: physiological data, deep learning, anomaly detection, fully convolutional neural networks, Encoder-Decoder architecture, AutoEncoder, Echocardiogram

Introduction

With the development of artificial intelligence, AutoEncoder Neural Networks have become an effective algorithm for the anomaly detection in various applications [1]. A variant of neural network trained using unsupervised learning is called an AE. Learning a comprehended form of the input data is achieved by the two primary components of an AE called as the encoder and the decoder [2]. The input is compressed into a latent-space representation by this network component. The input data is encoded in a reduced dimension as an encoded (compressed) representation. From the encoded representation, the decoder portion seeks to recover the input data. It attempts to produce an output that closely resembles the original input. In many of the application it was observed that AE's are effective at learning the distribution of the input data since they are trained to reduce reconstruction error [3].

In medical field, anomaly detection using machine learning is a significant problem that has been thoroughly researched. The emergence of deep learning has transformed the applications of machine learning, and its popularity has spread to the field of medical anomaly detection. The most typical diagnosis sequence for cardiac diseases starts with an ECG, then moves on to an echocardiography [4]. ECG is instant test to assess the heartbeat and it captures the electrical impulses produced by the heart. It can be used to determine arrhythmias, or abnormal heartbeats, and heart attacks.

ECGs can be collected easily using a wearable device which is a vital healthcare data for detecting and predicting the patients' health status. Most existing ECG analysis follows a feature extraction method that relies on rule-based approaches. It is difficult to manually define all ECG features. Many studies have attempted to develop a model that predicts a patient's disease or health state via electrocardiography; however, feature extraction is required to reduce the high-dimensional data included in ECGs. Various feature extraction methods have been proposed for ECG analysis, including morphological and temporal approaches [5]. An unsupervised feature learning method using AE architecture is proposed in this research.

Recently, there have been various attempts to use deep learning algorithms for extracting features from high-dimensional data. Supervised and unsupervised learning are the two methods used in deep learning. Feature extraction is used in supervised learning to categorize data based on a matched label. Every ECG feature in supervised feature learning needs to be labeled, which is time consuming and prone to error. As feature learning is task-specific, it is challenging to apply features extracted for supervised learning to different tasks. Neural networks which follow unsupervised learning techniques were used for the analysis of ECGs in order to overcome the challenges and drawbacks of supervised feature learning [6] and accurately represent the general features of ECGs.

In AE the original data and the reconstructed data were represented as points in an n-dimensional coordinate space for the purpose of anomaly detection, and the distance between the two points represents the difference between the given input and generated output which is referred as reconstruction error. When anomalous data is given as input, compression and restoration cannot be done efficiently since the AE used for anomaly detection is trained to recover normal data properly. As a result, the reconstruction error of the anomaly data is high. The data is classified as anomalous if the reconstruction error exceeds a predefined threshold; otherwise, it is classified as normal [7]. Depending on the unique characteristics of the input data, thresholds are defined in different ways. The experimental results presented in this paper compare the performance of the standard AE architecture with VAE and ensemble of AE's using bagging method. To lower prediction variance and generalization error, ensemble learning integrates predictions from several AE models.

Literature Review

Adopting a suitable anomaly detection technique to analyze physiological data in computer aided diagnosis can help to detect deviant patterns and help physician in early and accurate detection of abnormality [8]. In medical context normal data samples collected from healthy individuals are available in large volume but collecting data from patients having abnormal conditions is not easy. The scope and characteristics of anomaly can vary from one medical condition to other and few of the abnormal conditions are not observed ye. The following are a few challenges in modeling anomalies; less frequent, and distinct from each other. Modeling classifiers using such imbalanced data can lead to biased results and causes the model to predict majority classes as the result [9]. Test samples belonging to minority classes will be considered as noise and are ignored. Supervised classification approaches can be used when equal distributions of data samples are available for each class. Supervised classifiers are capable of modeling observed classes and they are incapable of representing unobserved anomalies. An artificial neural network type known as an AE can learn intricate correlations between the features in the data by training it in an unsupervised way using raw or processed data. AE, or its variants can learn latent representation by reconstructing the input and using an encoder and decoder architecture. Any appreciable variation in the reconstruction error during testing may be interpreted as unusual behavior or an occurrence. One benefit of AEs is their ability to process several kinds of biomedical data, such as CT, X-ray, ECG, and EEG [10].

Training an AE or one of its variants—such as convolution, sequential, or spatiotemporal—on a large amount of publicly available normal data and classifying significant deviations as anomalies based on reconstruction error or other metrics is a standard method for detecting anomalies. In many of the existing literature a single AE might not generalize well with noisy inputs and might be susceptible to the quality and preparation of the data [10, 11]. In few of the approaches to strengthen the anomaly detection process an ensemble of autoencoders was employed with an aim to integrate different autoencoders to arrive at a final decision. In this approach more than one model can be trained (individually or simultaneously), and the results are integrated to determine the ultimate output. To train multiple autoencoders, the different models could represent the multimodality or numerous viewpoints of the same data (raw video and optical flow). However, the production of unique datasets supplied to different AEs and the aggregation of individual AEs' results represent the main challenge in building an optimal Ensemble of AutoEncoder (EoAE). Compared to single AE, EoAE yields robust and accurate anomaly detection. The autoencoder's ability to reconstruct the input with less reconstruction error indicates that AE has performed well in generalizing the observed data. An unseen sample may be regarded as anomalous if it has a high reconstruction error, indicating a considerable deviation from the observed samples. AEs's capability to recognize these out-of-distribution samples serves as the foundation for anomaly detection.

A variational autoencoder does not learn any deterministic mapping from the input data to the encoded representation; rather, it learns the probability distribution of parameters [12]. This kind of AE is referred to as a generative model since it attempts to produce new data samples from the distribution sample. In contrast to basic AE, where output is mapped to input via bottleneck/latent space, in this instance the encoder representation's mean and standard deviation serve as the bottleneck, and sample latent space is used instead of the actual bottleneck. In few of the literature

In few of the literatures to lower prediction variance and generalization error, ensemble learning was adopted which integrates predictions from several AE models. By merging various and precise perspectives on normal data, an Ensemble of AEs (EoAEs) can function as a reliable method for anomaly detection in biomedical data. Compared to a single encoder, an EoAE can offer better detection; nevertheless, its effectiveness can rely on a number of elements, such as the variety of the generated data, the precision of each AE, and the combination of their results [15]. Utilizing the reconstruction probability derived from the variational autoencoder, an anomaly detection technique was suggested in [16]. Probabilistic metric that accounts for the variability of the variable distribution is the reconstruction probability. It was demonstrated in [17] that the reconstruction error, when used in conjunction with the backpropagated KL-term, can consistently increase performance for pixel-wise anomaly identification. This is comparable to popular VAE-based anomaly detection techniques that use the ELBO L, or the combination of the reconstruction and KL terms, as a proxy score.

Architecture of Variational AutoEncoder

Variational AutoEncoders (VAEs) models are specifically created to capture the underlying probability distribution of a given dataset and produce new samples using an encoder-decoder mechanism as part of their architecture [13]. The input data is transformed into a latent form by the encoder, and the decoder uses this latent representation to reconstruct the original data. The VAE can understand the underlying data distribution and produce new samples that follow the same distribution since it is trained to minimize the differences between the original and reconstructed data.

VAEs can be viewed as probabilistic rather than deterministic, in contrast to typical AE [14]. Given an input sample x_i, the compressed representation of x_i, z_i is created at random. Each high-dimensional datapoint x_i, originates from a lower dimensional data point z_i, according to the probabilistic VAE model. This generating process maps z_i to the distribution D's parameters that are needed to sample x_i using a neural network. VI allows us to deduce the model's parameters and latent variables. It is resolved by applying a neural network to map each x_i into parameters of the variational posterior distribution q needed to sample zi. The variational posterior $q_\phi(z|x)$ is combined with the distribution $p_\theta(x|z)$ that generates data. Given a data point x, z can be sampled from the variational posterior distribution $z \sim q_\phi(z|x)$ then a new data point x' is generated from $p(x|z)$ (mathematically it can be inferred as $x' \sim p_\theta(x|z)$). The loss function of the VAE is expressed mathematically as in Eq. 1;

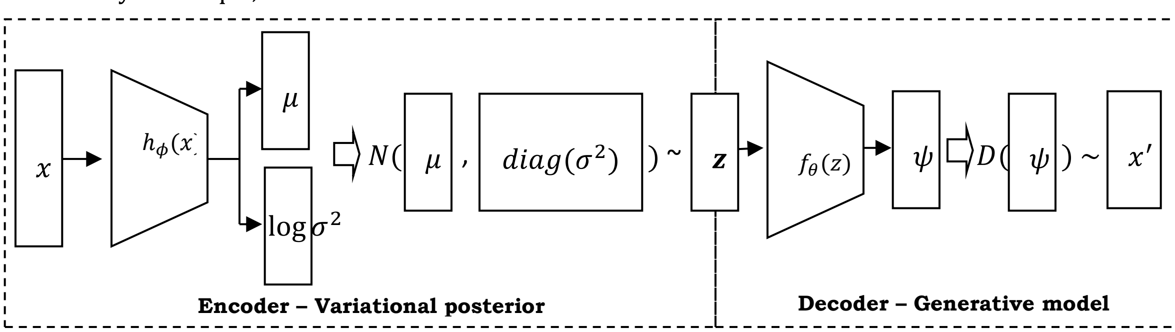

Fig. Schematic view of Variational AutoEncoder Reconstruction Process

$$loss_{VAE}(\phi,\theta) = -\sum_{i=1}^{n} E_{z_i \sim q_\phi(z_i|x_i)}[log\, p_\theta(x_i|z_i)] - KL(q_\phi(z_i|x_i)||p(z_i)) \qquad \text{Eq. 1}$$

$log\ p_\theta(x_i|z_i)$ represents the reconstruction error and it pushes the model to reconstruct the original x_i from its compressed representation z_i. This is achieve when $p_\theta(x_i|z_i)$ follows a normal distribution i.e. $p_\theta(x_i|z_i) := N(f_\theta(z_i), \sigma_{decoder}\boldsymbol{I})$ where $\sigma_{decoder}$ describes the volume of Gaussian noise $f_\theta(z_i)$. VAE minimizes the squared error between the decoded output x_i' and the actual input x_i. Using the analytical form of $log\ p_\theta(x_i|z_i)$ let us rewrite the Eq. 1 as follows;

$$loss_{VAE}(\phi,\theta) = \sum_{i=1}^{n} E_{z_i \sim q_\phi(z_i|x_i)}\left[\log\frac{1}{\sqrt{2\pi\sigma^2_{decoder}}} - \frac{\|x_i - f_\theta(z_i)\|_2^2}{2\sigma^2_{decoder}}\right] - KL(q_\phi(z_i|x_i)||p(z_i)) \qquad \text{Eq. 2}$$

On comparing the Eq. 2 with the loss function of the AE where the squared error is minimized as follows;

$$loss_{AE} := \frac{1}{n}\sum_{i=1}^{n}\left\|x_i - f_\theta(h_\phi(x_i))\right\|_2^2 \qquad \text{Eq. 3}$$

where h_ϕ is the neural network for encoding and $f_\theta$is the neural network which does the decoding. Both the general autoencoder and its variant VAE attempts to minimize the squared error between the reconstructed output x_i' and the input data point x_i. The VAE has a KL-divergence term with the opposite sign to the reconstruction loss term which is not present in a standard autoencoder. KL-divergence term pushes the model to synthesize the latent values from $q_\phi(z_i|x_i)$ which follow the (normal distribution) prior distribution $p(z_i)$. KL term acts as a regularization term on the reconstruction error. The objective of the model is to reconstruct every sample input x_i but also it makes sure that the latent representation z_i follows normal distribution.

Dataset Description

The patient's ECG readings are collected to construct this dataset. A patient's entire ECG is represented by each row in the dataset. There are 140 data points (readings) in each ECG. An individual patient's ECG data point can be found in columns 0-139. The marker that indicates whether an ECG is abnormal or normal is a categorical variable that has two possible values: 0 and 1. The waveform representation of the data samples is presented in Fig. 1. There are 4998 ECG samples including 2079 norma1 and 2919 anomalous sample data.

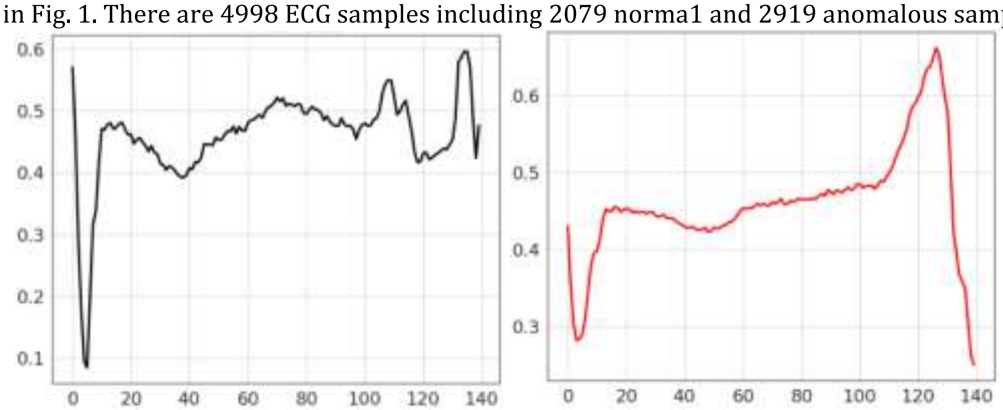

Fig. 1 Waveform representation of Normal and Anomalous ECG data

Experiments and Results

This section outlines the tests that were carried out to investigate the behavior of the AE and its variants. The ECG data points are initially normalized using min-max scaling method which transforms each feature by scaling them to a range between 0 and 1. The transformation of each feature can be represented mathematically as follows;

$$x_{std} = \frac{x - \min(x)}{\max(x) - \min(x)}$$
$$x_{scaled} = x_{std} * (\max - \min) + \min \qquad \text{Eq. 4}$$

The preprocessed data are used for training the training the VAE and AE models. To improve the performance of the AE model, an ensemble AE model was trained following bagging method. The performance of the various AE models were analyzed using different evaluation metrics including precision, recall, F1-score, Area Under Curve (AUC). An optimal threshold value has to be selected to fix the decision boundary that classifies normal and abnormal ECG samples. From the results presented in various literatures it is observed that the output from the AE model has prediction uncertainty. Various statistical approaches including maximum likelihood estimation and upper quartile has been studied for fixing the threshold. This paper uses a kernel density estimation (KDE) method to derive the probability density function (PDF) of the reconstruction error of the normal samples. The confidence interval of the normal data was also estimated by KDE method which is a non-parametric approach that estimates the PDF of samples following an unknown probability distribution. This approach does not require any prior knowledge about the data and any assumption. Let us consider that $R = \{r_1, r_2, \ldots, r_m\}$ is the sample derived from a probability distribution; H is a smoothing parameter and the kernel function is represented as $K(.)$ that satisfies the following condition as stated in Eq. 5;

$$K(r) > 0, \text{and} \int_{-\infty}^{\infty} K(r)d(r) = 1 \qquad \text{Eq. 5}$$

The estimation of the density function at a point r can be derived by ρ.

$$\rho(x) = \frac{1}{m\mathbf{H}} \sum_{i=1}^{m} K(\frac{X_i - x}{H}) \qquad \text{Eq. 6}$$

Based on the density function estimated by the KDE method the threshold limits $[\tau_l, \tau_u]$ are determined. τ_u is fixed based on the upper bound of the 90% of the confidence interval calculated from the probability density function.

$$\zeta(x) = \begin{cases} 0 \rightarrow \text{Normal when } \rho(x) \leq \tau_u \\ 1 \rightarrow \text{Anomaly when } \rho(x) > \tau_u \end{cases} \qquad \text{Eq. 7}$$

A higher likelihood of finding an abnormality is indicated by higher reconstruction error levels. VAE has the ability to learn important patterns and features in the raw input data which are compactly and meaningfully encoded in the latent space. Relative to the raw input features or the reconstruction error from a typical AE, this learned representation generated from VAE may be more useful for anomaly identification.

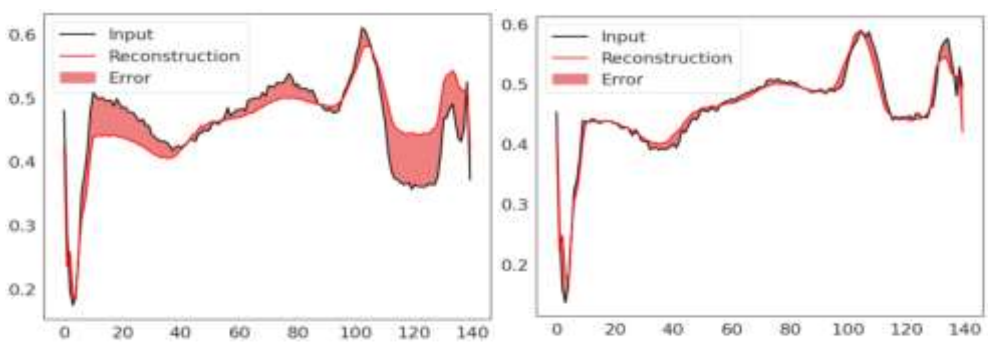

Fig. 2 Plot of VAE Reconstruction Error for test samples (normal data)

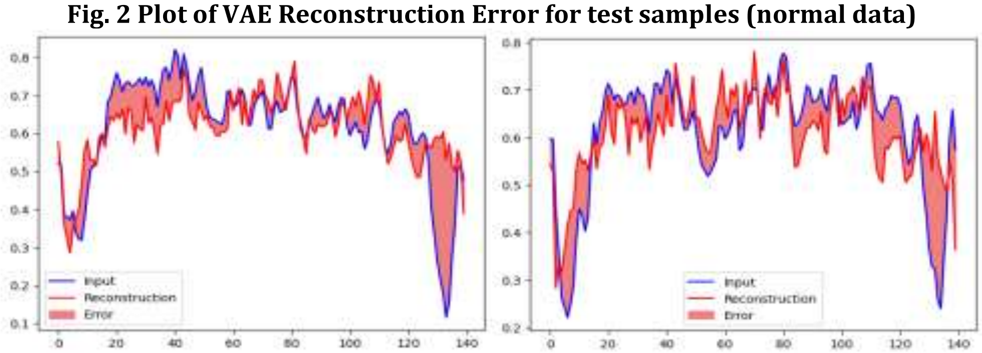

Fig. 3 Plot of AE Reconstruction Error for test samples (normal data)

Regularization is incorporated into VAEs through the loss function's Kullback-Leibler divergence term. This promotes a smooth and well-behaved latent space, which can enhance the model's capacity to generalize to new, normal samples and more accurately identify anomalies. The latent space uncertainty is modeled by VAEs, which might offer further information to distinguish between normal and abnormal data. Anomalous scoring also can be performed directly using the uncertainty estimations. VAEs outperform standard AE due to their generative modeling, probabilistic formulation, representation learning, and uncertainty modeling capabilities. Standard AE's are incapable of identifying complex relationships leading to high reconstruction error or poor feature extraction. Also they are prone to overfitting which is resolved in VAE by including KL-divergence term.

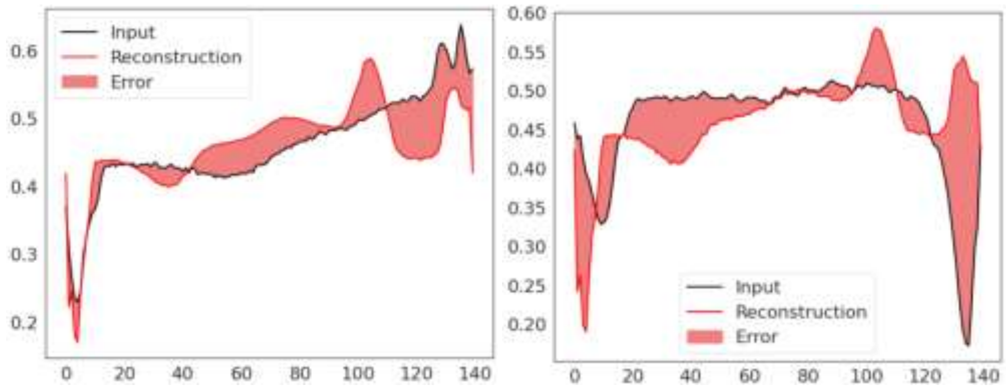

Fig. 4 Plot of VAE Reconstruction Error for test samples (abnormal data)

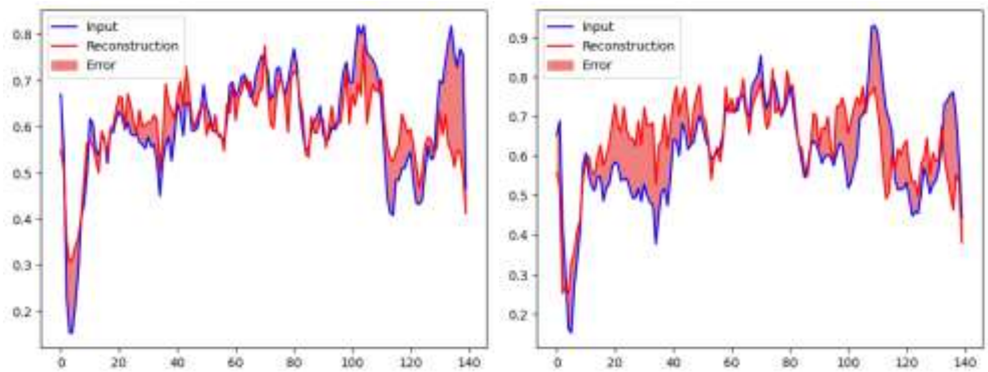

Fig. 5 Plot of AE Reconstruction Error for test samples (abnormal data)

By determining whether the reconstruction loss exceeds a predetermined threshold, anomalies are found. Test samples are categorized as abnormal if the reconstruction error exceeds the estimated threshold values. The mean error for normal examples in the training set is calculated.

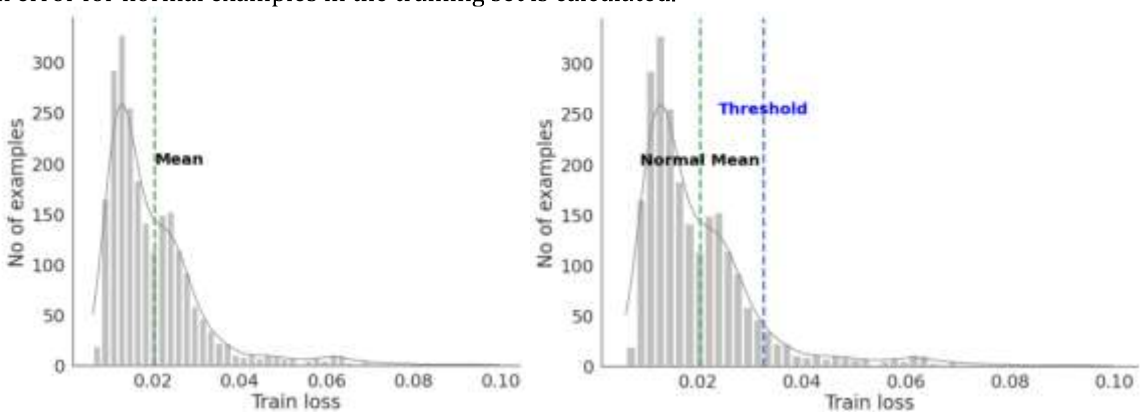

Fig. 6 Mean error for normal examples in the training set

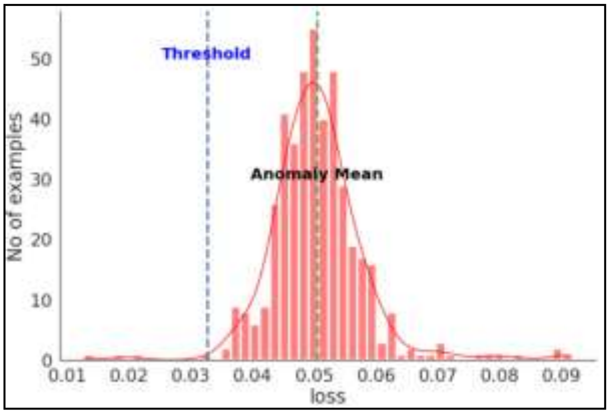

Fig. 7 Mean error for abnormal examples in the test set

As per the experimental results the majority of the test set's abnormal sample have reconstruction errors higher than the threshold. ECG signals are considered as abnormal if the reconstruction error exceeds the threshold (blue-dotted line). The results of the test samples were evaluated using confusion matrix. The Table 1 presents the results of VAE where the overall classification accuracy is 0.9440 with a precision value of 0.9452 and recall as 0.94. The estimated F1 score value was 0.9450 and AUC was 0.9435.

Table: 1 Result of anomaly detection on test samples (VAE)

	Precision	Recall	F1-Score	AUC
Normal	0.90	0.99	0.94	0.9490
Anomalous	0.99	0.89	0.93	0.9380

Table: 2 Result of anomaly detection on test samples (AE)

	Precision	Recall	F1-Score	AUC
Normal	0.86	0.92	0.88	0.8980
Anomalous	0.92	0.82	0.86	0.8926

Table: 3 Result of anomaly detection on test samples (EoAE)

	Precision	Recall	F1-Score	AUC
Normal	0.88	0.96	0.92	0.9250
Anomalous	0.96	0.88	0.92	0.9180

Discussion

VAEs exhibited better performance when compared to AE because of two main advantages they offer. VAEs provide for more precise control over the latent space's structure. The model's encoder will be pushed by the KL-divergence term to encode samples whose latent random variables are distributed similarly to the prior distribution. When learning low dimensional representations of data distributed according to more complex distributions, VAEs offer a systematic approach. VAE can be applied when ($p_\theta(x_i|z_i)$) is more complex. AE model learns training data too well—capturing noise and irregularities in the data. VAE model is encouraged to learn to accurately duplicate the most often seen characteristics by having it replicate the most prominent features in the training data under some of the limitations. The KL divergence term quantifies the degree to which the latent variable's learned distribution deviates from the previous distribution, which is typically taken to be a normal distribution. The VAE learns a more robust latent space and prevents overfitting with the aid of this regularization term. The pictorial representation of the normal and anomalous data samples (actual data and anomaly detected by VAE) in a reduced space using Principal Component Analysis is presented in Fig. 8.

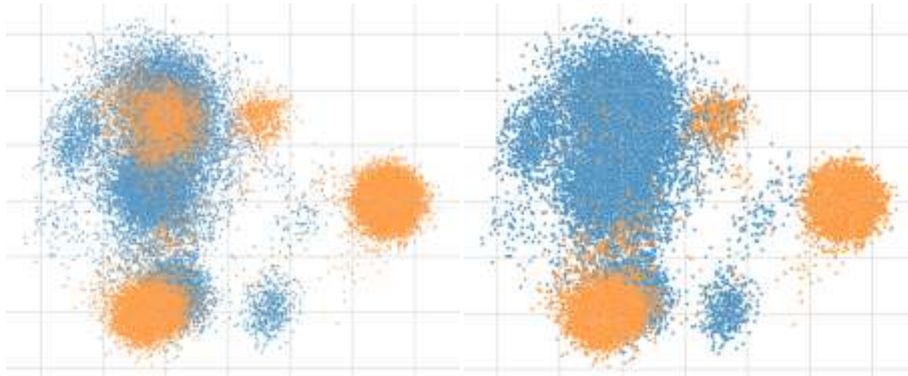

Fig. 8 Plot of Anomalous and Normal data points in reduced space (Orange – Anomalous; Blue – Normal | left - ground truth; right – VAE output)

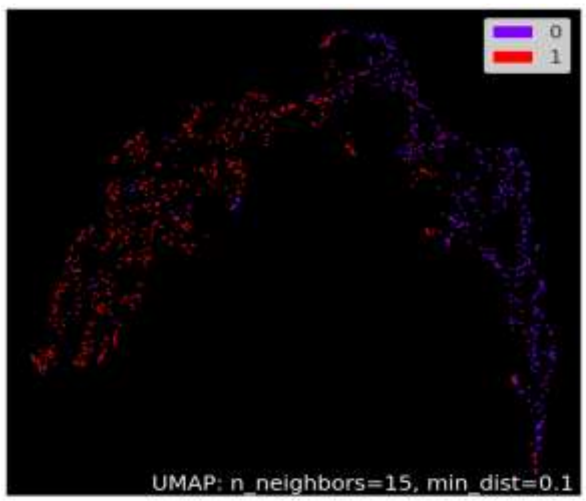

Fig. 9 Dimension reduction of the latent space

Dimension reduction of the latent space by VAE for both normal and pathological data is depicted in the Fig. 9. This latent space is the result of compressing an input of 140 dimensions into an 8-dimensional space through the process of Dimension Reduction. It is evident that in the latent space, normal and anomalous data are clearly divided. From a neural network perspective, this is the power of auto-encoders and a potential solution to the dimensional curse.

Conclusion

The problem of detecting anomalous ECG signals were formulated as a classification problem using unsupervised learning approach is followed. The efficiency of the unsupervised DL based variational autoencoder (VAE) model has been evaluated and compared with the performance of the autoencoder architecture and ensemble AutoEncoder. The efficiency of the classification approach is decided based on the threshold, hence careful selection of an optimal threshold value is essential. In case of anomaly detection using variants of auto-encoder network, recall is considered as an important evaluation metric when compared to precision. The regularization applied to the generated latent space is the main benefit of VAEs. Working with a smoother and more continuous vector space guarantees that comparable data points lie closer together, which improves the reliability of similarity measurements and can produce more consistent and accurate results.

References

1. Pang, Guansong, et al. "Deep learning for anomaly detection: A review." ACM computing surveys (CSUR) 54.2 (2021): 1-38.
2. Baldi, Pierre. "Autoencoders, unsupervised learning, and deep architectures." Proceedings of ICML workshop on unsupervised and transfer learning. JMLR Workshop and Conference Proceedings, 2012.

3. Asperti, Andrea, and Matteo Trentin. "Balancing reconstruction error and kullback-leibler divergence in variational autoencoders." Ieee Access 8 (2020): 199440-199448.
4. Suhail, M. Mohamed, and T. Abdul Razak. "Cardiac disease detection from ECG signal using discrete wavelet transform with machine learning method." Diabetes Research and Clinical Practice 187 (2022): 109852.
5. De Chazal, Philip, Maria O'Dwyer, and Richard B. Reilly. "Automatic classification of heartbeats using ECG morphology and heartbeat interval features." IEEE transactions on biomedical engineering 51.7 (2004): 1196-1206.
6. Cunningham, Pádraig, Matthieu Cord, and Sarah Jane Delany. "Supervised learning." Machine learning techniques for multimedia: case studies on organization and retrieval. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. 21-49.
7. Chen, Zhaomin, et al. "Autoencoder-based network anomaly detection." 2018 Wireless telecommunications symposium (WTS). IEEE, 2018.
8. Watrous, Raymond L. "Computer-aided auscultation of the heart: From anatomy and physiology to diagnostic decision support." 2006 International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2006.
9. Pang, Guansong, Longbing Cao, and Charu Aggarwal. "Deep learning for anomaly detection: Challenges, methods, and opportunities." Proceedings of the 14th ACM international conference on web search and data mining. 2021.
10. Nawaz, Ali, Shehroz S. Khan, and Amir Ahmad. "Ensemble of Autoencoders for Anomaly Detection in Biomedical Data: A Narrative Review." IEEE Access (2024).
11. Mujkić, Amar, et al. "Machine intelligence in biomedical data modeling, processing, and analysis." 2022 11th Mediterranean Conference on Embedded Computing (MECO). IEEE, 2022.
12. Doersch, Carl. "Tutorial on variational autoencoders." arXiv preprint arXiv:1606.05908 (2016).
13. Pinheiro Cinelli, Lucas, et al. "Variational autoencoder." Variational Methods for Machine Learning with Applications to Deep Networks. Cham: Springer International Publishing, 2021. 111-149.
14. An, Jinwon, and Sungzoon Cho. "Variational autoencoder based anomaly detection using reconstruction probability." Special lecture on IE 2.1 (2015): 1-18.
15. Nawaz, Ali, Shehroz S. Khan, and Amir Ahmad. "Ensemble of Autoencoders for Anomaly Detection in Biomedical Data: A Narrative Review." IEEE Access (2024).
16. An, Jinwon, and Sungzoon Cho. "Variational autoencoder based anomaly detection using reconstruction probability." Special lecture on IE 2.1 (2015): 1-18.
17. Zimmerer, David, et al. "Unsupervised anomaly localization using variational auto-encoders." Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part IV 22. Springer International Publishing, 2019.