FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

Acute Myocardial Infarction (AMI) Detection Using Multimodal Dataset, Optimized Variational Stacked Autoencoder (OVSAE) and Self-Attention Long Short-Term Memory (SALSTM) Classifier

Malathi R S[1], and Dr. Sudalaimuthu T[2]

[1]PHD Research Scholar; Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India; Email: rsmalathi50@gmail.com
[2]Professor and Head of the Department, Department of Computer Science and Engineering, Hindustan Institute of Technology and Science, Chennai, Tamil Nadu, India; Email: sudalaimuthut@gmail.com

***Correspondence:** *rsmalathi50@gmail.com*;

ABSTRACT– Acute Myocardial Infarction (AMI) is a vital public health concern, because it is the primary factor of death globally. Therefore, timely identification is crucial, especially in resource-constrained situations without centralized testing. (1) Background: Assessment, risk assessment, and treatment all depend on electrocardiograms (ECGs). ECG segments are artificially corrupted with various noise types (e.g., Gaussian noise, baseline wander) to create noisy training data.; (2) Methods: In this paper, signal denoising with an Optimized Variational Stacked Autoencoder (OVSAE) model which involves training a Neural Network (NN) to reconstruct clean ECGs from noisy versions, effectively learning to separate signal from noises and then decompressing the noise removed signals. OVSAE is introduced to adaptively remove noisy signals from ECG signals. VSAE learns hierarchical latent representations of clean signal structures through multiple nonlinear encoding layers. To get improved parameter initialization and optimum results, layer-wise prior training and modification are implemented. To enhance training stability and reconstruction accuracy, a layer-wise greedy pre-training strategy is adopted, followed by global fine-tuning of the entire network. Self-Attention Long Short-Term Memory (SALSTM) classifier is designed for AMI detection. To adaptively weight the significance of various temporal aspects and numerous ECG signals, the LSTM model incorporates a self-attention mechanism. A gating mechanism is added by LSTM, a gated network, to regulate the NN's information transfer; (3) Results: Clinical Parameters in Risk Stratification, and PTB-XL ECG diagnostic dataset includes of 18885 patients' 10-second clinical 12-lead ECGs, that of 21837.Furthermore, the results are quantified using measures like Mean Square Error (MSE), Mean Absolute Error (MAE), Structured Similarity Index (SSIM), and Peak Signal to Noise Ratio (PSNR). AMI results have been evaluated using metrics such as ERRor (ERR), ACCuracy (ACC), SPEcificity (SPE), and SENsitivity (SEN) against k-fold cross-validation.

Keywords: Acute myocardial infarction (AMI), Electrocardiogram (ECG), Optimized Variational Stacked Autoencoder (OVSAE), Self-Attention Long Short-Term Memory (SALSTM) classifier, and Deep Learning (DL)

ARTICLE INFORMATION

Author(s): Malathi R S and Dr. Sudalaimuthu T;
Received: 06/03/26; **Accepted**: 10/06/26; **Published**: 20/06/26;
E- ISSN: 2347-470X;
Paper Id: IJEER260110;
Citation: 10.37391/ijeer.140213
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-14/ijeer-140213.html

Publisher's Note: FOREX Publication stays neutral with regard to jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Cardiovascular diseases (CVDs) accounted for an estimated 18.6 million mortalities globally in 2024, and this figure is expected to increase to nearly 23.6 million by 2030. Approximately 85% of these fatalities are attributed to myocardial infarction (MI). For MI detection, a reliable diagnostic tool that is widely utilized is still the ECG [1]. AMI is a vital public health concern, because it is the primary factor of death globally. But the diagnostic methods and therapeutic interventions were greatly improved by this AMI. From acute myocardial ischemia, the necrosis of cardiomyocytes may result, and it is defined as AMI. The atherothrombotic processes are the main reasons for the MI infractions of Type1 and Type2. Distinct causes of myocardial ischemia are related to other types of AMI, may result in the death of myocyte [2]. The evidence of Acute Myocardial Destruction in medical situations that are consistent with myocardial ischemia, typically elevating cardiac biomarkers indicating myocardial necrosis leads to AMI. Non-ST-segment Elevation ACS (NSTE-ACS) and ST-segment Elevation MI (STEMI) are the 2 classes of Acute Coronary Syndrome (ACS), as per the outcomes of 12-lead ECG [3].

For risk assessment, and making therapeutic decisions, the analysis of ECG is very crucial. Early detection and prompt diagnosis is crucial for the ventricular arrhythmias, because it occurs in the beginning of the presentation hours, and it is the major reason for lots of early mortality rate. Immediate

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

reperfusion therapy is crucial If a patient having symptoms of myocardial ischemia and STEMI on ECG. This therapy will help in reducing cardiac morbidity and mortality of the patient [4]. Informative predictions are offered by continuous ECG surveillance. The evaluation of successful reperfusion and the detection of re-occlusion are facilitated by this surveillance [5]. The suspected cases from ECG evaluation support providing some vital initial diagnostic methods in both pre-hospital and in-hospital settings. Specific anatomical regions of the heart are matched to individual ECG. According to the anatomical position of myocardial cell death, the MI types are differentiated (*figure 1*) [6]. The 12-lead interpretation becomes challenging and time-sensitive when there are heterogeneous and complex ECG presentations in MI, this may increase the risk of diagnostic errors.

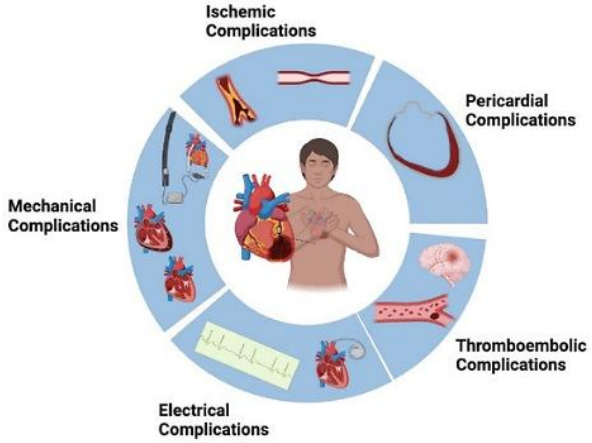

Figure 1. Complications of AMI

For risk assessment, and making therapeutic decisions, the analysis of ECG is very crucial. Early detection and prompt diagnosis is crucial for the ventricular arrhythmias, because it occurs in the beginning of the presentation hours, and it is the major reason for lots of early mortality rate. Immediate reperfusion therapy is crucial If a patient having symptoms of myocardial ischemia and STEMI on ECG. This therapy will help in reducing cardiac morbidity and mortality of the patient [4]. Informative predictions are offered by continuous ECG surveillance. The evaluation of successful reperfusion and the detection of re-occlusion are facilitated by this surveillance [5]. The suspected cases from ECG evaluation support providing some vital initial diagnostic methods in both pre-hospital and in-hospital settings. Specific anatomical regions of the heart are matched to individual ECG. Based on the anatomical position of MI, the MI types are differentiated (*figure 1*) [6]. The 12-lead interpretation becomes challenging and time-sensitive when there are heterogeneous and complex ECG presentations in MI, this may increase the risk of diagnostic errors. Transforming waveform images into digital signal sequences is the main emphasis of conventional ECG image processing methods. It executes geometric corrections and able to detect characteristic patterns. These conventional methods are vulnerable to noise and artifacts like printing distortions, inadequate skin–electrode contact, and external interference, all these impacts the accuracy and delivering low efficiency than experienced doctors like emergency medicine residents or cardiologists. So, these conventional methods are limited to use [7–8]. The following noise sources then affect the ECG data: electromyographic activity from muscle contractions, power-line interference (PLI), and baseline wander (BW).

The diagnostic accuracy of automated methods is greatly impacted in an adverse manner by this noise contamination, and it is the reason behind poor signal quality. Before the feature extraction process, the prior step of denoising takes place. Signal distortion may result from the application of conventional denoising methods. Clinically relevant attributes are also reduced by this implementation. Eliminating these random, avoidable, or artifactual components has become the objective of signal denoising for improving the signal-to-noise ratio (SNR) and maintaining critical data. Inspired by all these factors, a Deep Learning (DL)-based approach for denoising ECG signals is suggested.

The MI diagnosis mostly depends upon manual, expert-driven interpretation of ECG recordings. So, it is time-sensitive and has the probabilities to make error especially in extended assessments. Automated MI detection has gained more interest with the quick advancements of computer science and signal processing techniques. The DL and conventional Machine Learning (ML) models are utilized by sophisticated methods for the purpose of resolving inherent drawbacks of manual analysis [9-10]. Advanced DL methods are contributed by the recent advancements in this domain, and it may support in providing comprehensive ECG analysis, like applications in arrhythmia classification and AMI detection.

For ECG signal denoising, an optimized Variational Stacked Autoencoder (OVSAE) framework is introduced. Here, the NN is trained by the suggested model for reconstruction of ECG signals. The noisy counterparts in the ECG signals can be removed and reconstructed into a clean ECG signal by this suggested method. By learning this separation, OVSAE produces denoised and compressed representations of ECG signals. SALSTM classifier is introduced for Acute Myocardial Infarction (AMI) detection. In order to improve feature representation, the model can dynamically assign relevance to various temporal features and multiple ECG leads based on the self-attention mechanism. Then, risk stratification is supported by integrating the PTB-XL ECG diagnostic dataset, and clinical parameters. The effectiveness of the suggested method can be determined by using standard quantitative measures.

A deep encoder–decoder denoising architecture called LU-Net was created by Ali et al. [10] to reduce lung sound interference from both internal and external sources. Overlaying of already obtained obvious heart sound recordings with hidden lung sounds and ambient hospital noise developed two open-access heart sound (OAHS) noisy datasets for evaluating performance. Ambient noises that arise in naturally occurring and artificially polluted recordings were reduced by the suggested method, and it also supports enhancing SNR, based on the simulation outcomes. It is validated that the recommended framework has the potential to attain superior

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

average SNR in the presence of lung sounds and unidentified clinic ambient noise. With this superior SNR, the recommended method executes well than the standard U-Net framework. Thus, the classification accuracy is improved by the denoising method. This method was analyzed over multiple datasets with diverse noise levels, and heterogeneous noise patterns. From the outcomes, it is validated that this architecture attains robust efficiencies over multiple datasets. With the aim of improving the efficiency of computer-aided auscultation systems, a DL-based phonocardiogram (PCG) denoising technique is suggested in noisy and low-resource hospitals. The following metrics: Root MSE (RMSE), Percent Root Mean Square Difference (PRD), estimated SNR, and actual SNR can be utilized to determine the efficiency of denoising.

Using the inherent periodic structure of ECG waveforms, a DL-driven ECG denoising technique was suggested by Rasti-Meymandi and Ghaffari [11]. Here, 2-D signals are formed by aligning and aggregating individual cardiac cycles. Then, the Convolutional NN (CNN) receives those 2-D signals. The correlation among the cardiac cycles is used, and it supports the model in attaining reliable and effective ECG denoising. A novel Local/Non-local Cycle (LNC) observation module is integrated in this CNN framework for capturing those inter-cycle correlations. The efficiency of the model can be assessed using MIT-BIH Arrhythmia Database in terms of improvement in SNR (SNRimp) and RMSE.

To detect and locate MI utilising a Multi-Channel Dense Attention NN (MCDANN), a 12-lead ECG recording was utilised by Qiang et al. [12]. Twelve parallel channels are included in this MCDANN. Each channel is designed for extracting heartbeat-related attributes from an individual lead. To detect and localize MI, those extracted attributes are then fused. In every channel, the dense layers and Squeeze-and-Excitation (SE) modules are mostly integrated. The feature data reusage in the lead is enhanced by the dense layers. Then, SE block is responsible for suppressing less critical signals in classifying ECG. This MCDANN framework is assessed using the PTB Diagnostic ECG Database. From the outcomes of the simulation, this suggested model can deliver high ACC and F1-scores in detecting and localizing MI for both intra-patient and inter-patient evaluation situations. When compared to other standard methods, this method attains significant performance in localizing inter-patient MI.

For classifying ECG signal, a Stacked Denoising Autoencoder (SDAE) was suggested by Gunawan et al., [13]. For the purpose of improving arrhythmia diagnostic accuracy, this method is used. From the input signals, noises are eliminated by this model, and learn robust, high-level features. This may enhance efficiency in classification. Here, the data pre-processing, SDAE feature learning and DNN was integrated in this recommended method, and the MIT-BIH Arrhythmia Database is utilized for analysis. A superior and effective ECG identification was attained by the integration in this method.
To improve the ECG signal, a Fully Convolutional Network (FCN)-based Denoising Autoencoder (DAE) was presented by Chiang et al. [14]. The data was compressed effectively by reducing the dimension of the input signal to 32 times than the original. This method is assessed using MIT-BIH Arrhythmia Database, it contains both ECG signals and noise signals. The efficiency in denoising is determined by using the metrics like SNRimp, RMSE, and PRD. Noisy ECG signals with various SNR levels are used to assess the experiment. From the simulation outcomes, it is clear that the suggested FCN-based method executes well than the traditional CNN-based denoising models. In conclusion, this framework is a strong tool in suppressing high noise, and attains low RMSE and PRD values, and high SNRimp.

A deep convolutional encoder–decoder network (DCNN-ED) with symmetric skip connections is suggested by Fotiadou and Vullings [15]. This may help in learning end-to-end transformations from noise-contaminated fetal ECG data and cleaning its counterparts. From the data of the simulation, for fetal ECG signals, an average SNR gain of 9.5 dB is attained. Utilizing a vast array of actual signals, this method is then simulated. From the outcomes, it is proved that this method has the potential to deliver superior quality improvement in fetal ECG signals. The multi-channel inputs are utilized by this method. The recommended method does not need any past data regarding the pulse location or noise power spectra, and it can maintain beat-to-beat morphological variations.

A novel Multi-Lead Diagnostic Attention-based Recurrent NN (MLDA-RNN) was suggested by Prabhakararao and Dandapat [16]. The three MI severity phases from HC participants can be automatically diagnosed using this method. The 12-lead ECG was analyzed by this framework. From all lead, the multi-scale temporal dependencies are captured, and it may help in improving classification. To obtain lead-specific discriminative features, the intra-lead attention module receives those features. The MLDA-RNN model is simulated on the PTB Diagnostic ECG Database and the STAFF III Dataset. High accuracy and superior class-specific detection performance was attained by this model, as demonstrated by the simulation outcomes. The learned attention weights mostly correlate with the doctor's method of diagnosing the severity stages of MI. This recommended method also attains superior performance, and it is also proven to be effective for model interpretability. The following metrics: ACC, SPE, F1-score, weighted PRE, Recall, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC) were implemented for evaluating the model's efficiency.

To differentiate the normal individuals and patients in ten MI categories based on the anatomical site of infarction, a Densely Connected Convolutional Network (DenseNet) in conjunction with a traditional CNN was created by Jahmunah et al. [17]. From the PTB Diagnostic ECG Database, ECG signals are pre-processed prior to classification. Then, with the R peak detection algorithm, ECG beasts were extracted. The 2 frameworks receive the beats separately. For classification purposes, DenseNet is the effective and suitable model, because of its low computational complexity. Greater classification in precision is also attained by this DenseNet model, because of its feature reusability. Here, an advanced Class Activation Mapping (CAM) method named Grad-CAM

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

was applied, and it may contribute to offering the ECG leads and the areas of ECG waves, that have become influential for model's predictive decision-making over 11 classes. From the DenseNet and CNN models, the most activated lead is the lead V4, validated by analysis. Comprehensible insights regarding the decisions that are taken by model is offered in this study, so it is clinically vital. DenseNet and CNN models have the ability to be used for ECG triage of MI treatment in clinics and remote out-of-hospital situations if they receive necessary clinical acceptance. The efficiency of this framework is assessed with the following metrics: ACC, SPE, and SEN.

A ResNet –Transformer Cascaded Network (RTCN) was suggested by Chen et al. [18]. From the S-transform, the time-frequency of the ECG signal features are obtained, and it can be analyzed by this suggested model. To obtain temporal frequency characteristics from the ECG data in an adaptive manner, this S-transform is utilized. Conventional methods fail to detect dynamic variations in cardiac cycles, but these variations are effectively captured by the scalable Gaussian window and high phase revolution of this transform. For extracting multi-scale local features and long-range temporal dependencies, a method that integrates the Transformer attention mechanism with ResNet is used. Both the local features and global features are not effectively captured by those current DL models; it doesn't have such potential to detect features. To create high-fidelity ECG data for minority classes, the Denoising Diffusion Probabilistic Model (DDPM) is used. Thus, the inter-patient classification performance was also improved. The abnormal characteristics were the main emphasis of this framework, and it was verified by the Gradient-weighted CAM (Grad-CAM). Its medical comprehension was also validated. Its efficiency was computed by ACC, SEN, PRE, SPE, and F1-score.

To diagnose MI using ECG signals, an Adaptive RNN (ARNN) was created by Mahendran et al. [19]. From ECG signals, accurately detecting MI disease has become the main objective of this suggested work. Then, the multi-Notch filter was utilized for denoising ECG signal. Specified noise frequency range is also eliminated with this filter. For feature extraction, Discrete Wavelet Transform (DWT) is used. With wave filtering bank, this DWT has the potential to break down ECG signals to varied scales. The DL-based ARNN classifier is utilized for classifying the defected and normal ECG arrhythmic beats after specific QRS attributes are obtained. The simulation was conducted to train and test the data of the framework using the arrhythmia database from the Massachusetts Institute of Technology–Beth Israel Hospital (MIT-BIH). Based on the accuracies in classification, the recommended method is assessed. When comparing the recommended work with others, this recommended method executes well and attains superior classification accuracy than other models. SEN, SPE, ACC, Negative Predictive Value (NPV), and Positive Predictive Value (PPV) were used in the analysis.

In order to classify ECG signals into three classes: Normal Sinus Rhythm (NSR), Atrial Fibrillation (AF), and Non-AF (NAF), a hybrid framework combining the DWT with one-dimensional CNNs (1D-CNNs) was created by Nurmaini et al. [20]. Three open datasets and one dataset from Indonesia hospital were integrated to assess the model in simulation. From the outcomes, it is clear that, the 1D-CNN model attains high ACC, SEN, SPE, PRE, and F1-score for all the 3 classes, and also attains high generalization ability. It is also demonstrated from the outcomes that this 1D-CNN model executes well than the conventional models. Early detection and better treatment results are obtained by this method with patient self-monitoring and clinical AF diagnosis.

For automatic MI, a 1D-CNN framework with 15-lead ECG signal (12 standard + 3 frank leads) was suggested by Mirza et al., [21]. There are 5 stages in this methodology; they are signal normalization, DWT-based denoising and R-peak detection, ECG segmentation, classification using a 1D-CNN architecture, and performance evaluation with ACC, SEN, SPE, PRE, and F1-score. The simulation was executed on the Physikalisch-Technische Bundesanstalt (PTB) dataset to measure the efficiency of the 1D-CNN frameworks. The ECG signals from 52 normal persons and 148 affected persons diagnosed with MI are included in this dataset. From the outcomes, it is clear that the suggested method attains average ACC, SEN, SPE, PRE, and F1-scores, and detects MI cases more effectively.

For CVD classification, a novel method for ECG signal processing named A DNN-based framework was suggested by Khatar et al. [22]. For ECG signal denoising, a new combined numerical filter, the central element of the study is used. The Chebyshev, Butterworth, and Daubechies filtering techniques are fused in this filter. To effectively denoising ECG signal, this filter is optimized. Relevant features are maintained by this filter. The reliability and accuracy of the cardiac abnormality detection and classification particularly in arrhythmia diagnosis are greatly enhanced by the application of this filter in DNN. The metrics like AUC-ROC, F1-score, SEN, SPE, and ACC are utilized for assessing the overall ECG denoising and CVD classification system.

An automated, scalable system for MI identification and localisation with 12-lead ECG records was suggested by Pan et al. [23]. The suggested approach seeks to reduce cardiologists' workload. To effectively obtain and combine features across several leads, the suggested Multi-Task Channel Attention Network (MCA-Net) makes use of a residual-based channel attention method. The total efficiency of the framework is enhanced by using a multi-task learning architecture to take use of valuable and shared data among the localization and MI detection tasks. The MCA-Net approach can effectively support cardiologists in identifying and localizing myocardial infarction, contributing to earlier diagnosis and improved patient prognosis.

2. MATERIALS AND METHODS

Signal denoising with an OVSAE involves training a NN to reconstruct clean ECGs from noisy versions, effectively learning to separate signal from noises and then decompressing the noise removed signals. OVSAE is

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

introduced to adaptively remove noisy signals from ECG signals. VSAE uses several nonlinear encoding layers to learn hierarchical latent features of clean signal patterns.

Layer-wise pre-training and fine-tuning are used to get optimized parameter initialisation and optimal outcomes. AMI detection is the purpose of the SALSTM classifier. Finally, results are measured using the metrics like PSNR, MSE, Structured Similarity Index (SSIM), and MAE. Overall process of proposed model is shown in *figure 2*.

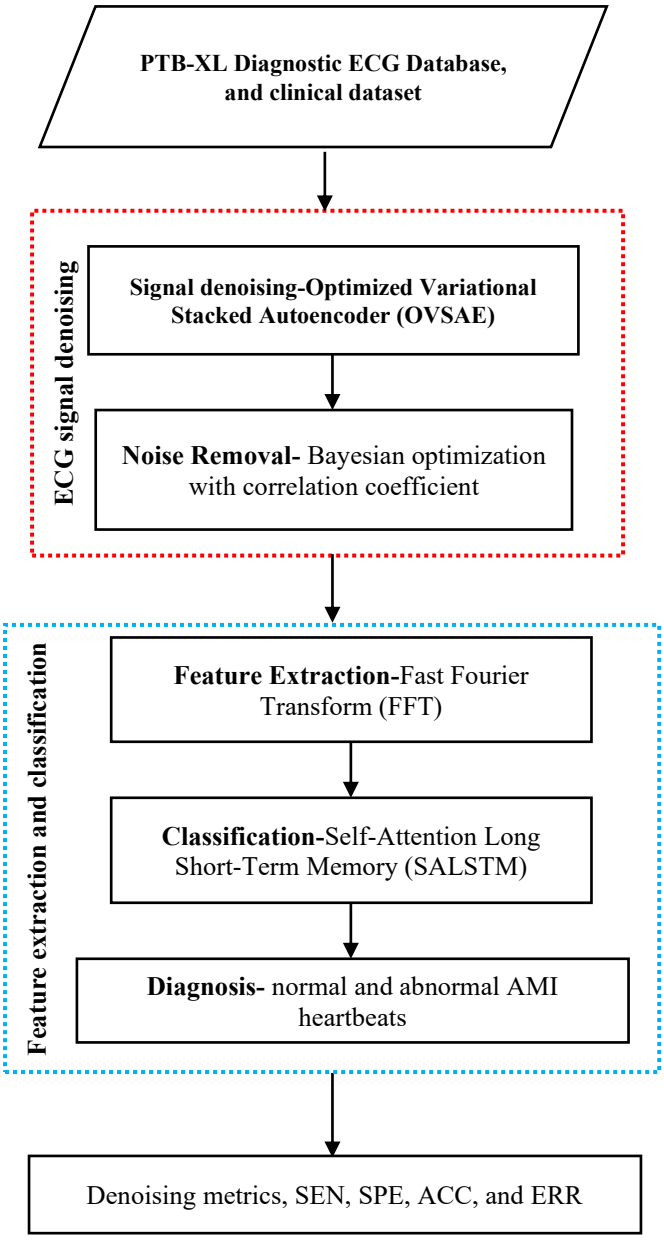

Figure 2. Overall Process of Proposed Model

2.1. Dataset

The following datasets has been used for heart attack detection. Heart Attack Dataset: The https://www.kaggle.com/datasets/sukhmandeepsinghbrar/heart-attack-dataset dataset. Age, gender, heart rate, blood pressure (BP) (systolic and diastolic), blood sugar, CK-MB, and troponin are all included in this dataset, and the results might be either positive or negative.

The male and female are assigned to 1 and 0, respectively, in the gender column of the data, which is normalized. If over 120, if it is not set to 1, the glucose column is fixed to 0. The result is a positive number one and a negative number zero. PTB-XL ECG database: https://www.kaggle.com/datasets/khyeh0719/ptb-xl-dataset/data

PTB-XL ECG diagnostic database a widely used dataset in the field of cardiology. Myocardial infarction, a critical medical condition, requires prompt diagnosis and treatment. Machine learning algorithms have the ability to support clinicians in accurately identifying myocardial infarction cases from ECG data. Additionally, 21,837 10-second clinical 12-lead ECG recordings from 18,885 individuals make up the PTB-XL ECG dataset.

A maximum of two cardiologists annotated the raw waveform data, and each record may have been given more than one ECG statement. Based on the SCP-ECG standard, 71 various ECG reports (including detection, morphological, and rhythm-related categories) are presented. Owing to its comprehensive annotation scheme, the dataset functions as a significant asset for the development and assessment automated ECG interpretation systems.

2.2. Optimized Variational Stacked Autoencoder (OVSAE) Based Signal Denoising

An ECG monitors the electrical activity of the heart to assess a range of cardiac conditions, despite the fact that noise frequently distorts the recorded data. Denoising of ECG signals constitutes a fundamental preprocessing stage that reduces unwanted disturbances while enhancing characteristic waveform components. This process targets the removal of baseline deviations, power line interference (PLI), and other artifacts, and electromyography noise, thereby facilitating reliable detection of cardiac abnormalities. Effective ECG denoising seeks to minimize noise contamination while preserving the diagnostically significant P–QRS–T complexes.

2.2.1. Baseline Wander (BW): 0.05–0.5 Hz low-frequency drift brought on by breathing and movement of the body.

2.2.2. PLI: Sinusoidal noise in a narrowband at 50/60 Hz. The incorrect diagnosis results from P-wave distortions.

2.2.3. Muscle Artifacts (MA): High-frequency noises (20–150 Hz) are generated by muscular activity.

The noisy ECG signal is expressed by *equation (1)*,

$$x_{noisy}(t) = x_{clean}(t) + n_{BW}(t) + n_{PLI}(t) + n_{MA}(t) \quad (1)$$

The signals are segmented, normalized, and for the noise reduction training. As illustrated in Figure 3, the encoder applies a nonlinear transformation to compress the decoder

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

uses this compressed space to recreate the ECG signal after transforming the input signal into a latent representation. The output layer is trained to generate vectors that closely approximate, but do not exactly replicate, the original input. In this stage, the hidden layer outputs represent features automatically learned from the ECG signal through unsupervised training. By reducing reconstruction error, the Autoencoder (AE) maintains the structural elements of the *P* wave, QRS complex, and *T* wave.

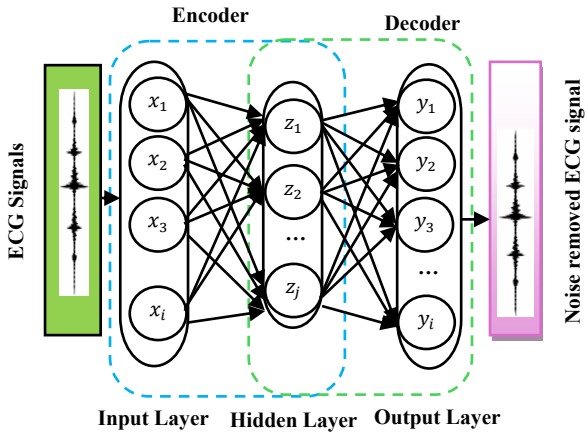

Figure 3. Design of an AE

Equation (2) provides a description of the encoding process,

$$Z = \sigma_1(W_1 \cdot X + b_1) \quad (2)$$

where the input and hidden layers' respective vectors are denoted by X and Z.$W_1$and $b_1$are, respectively, the basis vector and weight matrix.$\sigma_1$corresponds to the encoder's activation function. *Equation (3)* describes the decoding procedure,

$$Y = \sigma_2(W_2 \cdot X + b_2) \quad (3)$$

The bias vector is denoted as b_2. The weight matrix is denoted as W_2. The output vector is denoted as Y. The activation function of the decoder is denoted asσ_2. Due to its powerful generative modeling and representation-learning capabilities, a Variational AE (VAE) has been frequently used for ECG signal denoising [24]. *Figure 4*, by sampling from the latent variable z's distribution, VAE produces the observed ECG signal *x*.

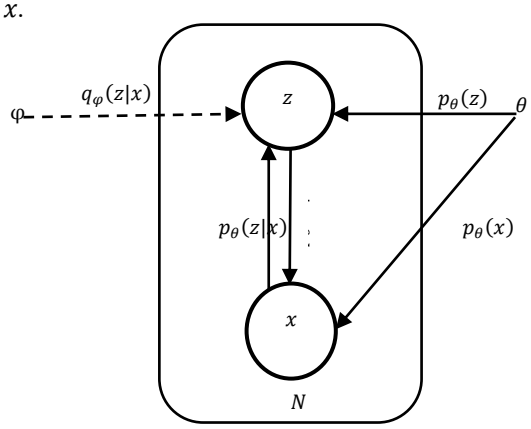

Figure 4. Graph Model of VAE

The prior probability of the z is denoted as $p_\theta(z)$, and it represents the assumed clean latent structure of ECG signals; this represents z's initial distribution. The distribution of the noisy ECG signal x that must be produced is $p_\theta(x).p_\theta(z|x)$ is another name for the encoder, which is the posterior distribution of z. Considering the fact that$p_\theta(z|x)$ that reconstructs the ECG waveform conditioned on z, the variationalapproximation$q_\varphi(z|x)$is often used to replace it. By optimizing the parameters θ and *φ*to minimize the divergence between$q_\varphi(z|x)$and$p_\theta(z|x)$, the VAE learns a smooth and noise-invariant latent representation of clean ECG signals. When trained with noisy ECG inputs and clean ECG targets, the model preserves the most important morphological features of the ECG waveform while successfully suppressing baseline variation, power-line interference, and muscle deformations [25].

In VAE, Gaussian distribution is typically denoted as $q_\varphi(z|x)$, Equation (4) [26-27] describes the Kullback-Leibler (KL) divergence between $p_\theta(z|x)$ and $q_\varphi(z|x)$,

$$D_{KL}[q_\varphi(z|x)\|p_\theta(z|x)] = E_{q_\varphi(z|x)}[\log q_\varphi(z|x) - \log p_\theta(z|x)] \quad (4)$$

Equation (4) can be expressed using the Bayes rule in *equation (5)*.

$$D_{KL}[q_\varphi(z|x)\|p_\theta(z|x)] = E_{q_\varphi(z|x)}[\log q_\varphi(z|x) - \log p_\theta(z|x) - \log p_\theta(z)] + \log p_\theta(x) \quad (5)$$

This is the formula for KL divergence, which is non-negative in *equation (6),*

$$\log p_\theta(x) \geq L(\theta, \varphi; x) \quad (6)$$

$$(\theta, \varphi; x) = -E_{q_\varphi(z|x)}[\log q_\varphi(z|x) - \log p_\theta(z|x) - \log p_\theta(z)] = -D_{KL}[q_\varphi(z|x)\|p_\theta(z)] + E_{q_\varphi(z|x)}[\log p_\theta(x|z)] \quad (7)$$

In *equation (7)*, $L(\theta, \varphi; x)$is referred to as the low boundaries of variance of$\log p_\theta(x)$it also serves as VAE's loss function. The phrase for regularization is the first one. The inaccuracy in reconstruction is shown by the second phrase. So$p_\theta(z) \sim N(0, I)$, $q_\varphi(z|x) \sim N(\mu, \sigma^2)$, the RHS of the 1st term on *equation (7)* is computed in *equation (8):*

$$-D_{KL}[q_\varphi(z|x)\|p_\theta(z)] = \frac{1}{2}\sum_{j=1}^{d}\left(1 + \log(\sigma_j)^2 - (\mu_j)^2 - (\sigma_j)^2\right) \quad (8)$$

Here, the distribution's dimensionality is denoted by *d*. *Equation (9)* provides calculation of RHS of 2nd term in *equation (7),*

$$E_{q_\varphi(z|x)}[\log p_\theta(x|z)] = \frac{1}{L}\sum_{j=1}^{d} \log p_\theta(x|z^{(l)}) \quad (9)$$

where L is chosen to be 1 by *equation (10)*,

$$z^{(l)} = \mu + \sigma \odot \varepsilon^{(l)}, \varepsilon^{(l)} \sim \mathcal{N}(0, I) \quad (10)$$

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

VAE loss in Equation (7) can be optimized using backpropagation despite the stochastic sampling process. The model effectively suppresses baseline wander, muscular artifacts and power-line interference while maintaining the ECG signal's structural characteristics, making it suitable for robust ECG denoising.

2.2.4. OVSAE: Typically, an SAE model is constructed by stacking numerous AEs to improve high-level signal denoising. SAE to directly learn discriminative features from raw signals, such as vibration or ECG signals. SAE performs signal denoising only through deterministic nonlinear transformations and therefore struggles to model the global structure of the latent signal space. This limitation can be effectively addressed by introducing a VAE. The VSAE is a hybrid neural architecture that integrates a VAE within a SAE, thereby leveraging the strengths of both models. As illustrated in Fig 5, 2 autoencoders are first pre-trained sequentially. Subsequently, AE1, AE2, and the VAE are combined to construct the OVSAE framework. Parameters of the framework are further refined through Bayesian optimization. AE1 and AE2 first construct hierarchical feature representations of the input signals within the VSAE structure. The dispersion of the latent feature space is then characterized using a VAE.

Specially, two encoders are utilized to estimate the mean vector (μ) and standard deviation vector σ corresponding to multiple Gaussian distributions. Samples are then drawn from N(0, I), a typical normal distribution, and the latent representation z is generated according to Equation (10), which formalizes the parameterization procedure. At last, the obtained features are reconstructed into signals that closely resemble the original ECG inputs by the decoders, which are made up of several fully connected (FC) layers.

The OVSAE model captures the most significant signal characteristics, and the correlation coefficient determines the variation between the initial ECG signal and the denoised output to assess reconstruction quality.

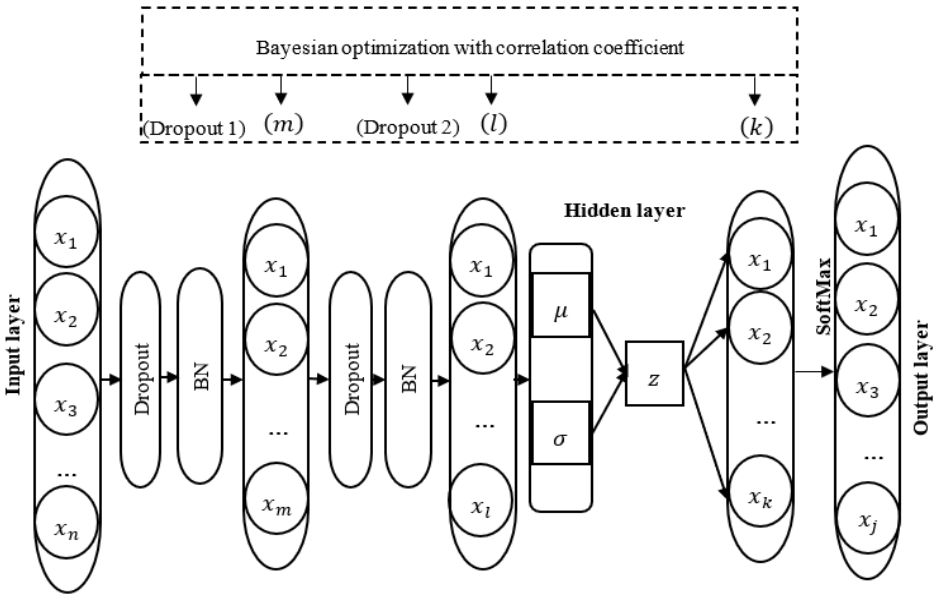

Figure 5. Optimized Variational Stacked Autoencoder (OVSAE)

For ECG denoising, the decoder component of the VSAE is eliminated and the FC layer and softmax layer are added in *figure 5*. Dropout layers are introduced, and Batch Normalization (BN) is fine-tuned to complete model training. Five key hyperparameters are considered: Dropout1, Dropout2, and the hidden layer sizes *m*, *l*, and *k*. The initial parameter intervals are defined based on prior experimental results and domain knowledge. Subsequently, the search space and harmony memory are constructed in accordance with these predefined ranges. Bayesian optimization, guided by the correlation coefficient, is then used to adaptively estimating the optimal values of the five hyperparameters. The OVSAE approach demonstrates strong diagnostic accuracy and robust generalization performance across diverse operating conditions. The correlation coefficient is measured between$x$and x' is calculated by *equation (11)*,

$$r(x,x') = \frac{cov(x,x')}{\sqrt{Var(x)Var(x')}} \quad (11)$$

The covariance of x and x' is represented as $cov(x,x')$.The variance of$x$is denoted as Var(x). The variance of x' is denoted as Var(x'). After denoising, the latent features obtained from the OVSAE encoder are used as input to the SALSTM network for temporal modeling and AMI classification.

2.3. Fast Fourier Transform (FFT) based Feature Extraction

The FFT method uses a recursive divide-and-conquer strategy to deal with time, based on the Discrete Fourier Transform (DFT) [28]. FFT splits the input by even and odd indices using a recursive divide-and-conquer strategy in *equation (12)*,

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

$$X_k = \sum_{j=0}^{\frac{n}{2}-1} x_{2j}.\omega_n^{2j.k} + \sum_{j=0}^{\frac{n}{2}-1} x_{2j+1}.\omega_n^{(2j+1).k} = \underbrace{\sum_{j=0}^{\frac{n}{2}-1} x_{2j}.\omega_{n/2}^{j.k}}_{DFT(x_0,x_2\ldots)_k mod n/2} + \omega_n^k \underbrace{.\sum_{j=0}^{\frac{n}{2}-1} x_{2j+1}.\omega_{n/2}^{j.k}}_{DFT(x_1,x_3\ldots)_k mod n/2} \quad (12)$$

An equivalent FFT algorithm is performed for smaller sequences, the FFT is determined recursively, and the results are combined [28]. This method starts with the input being permuted, then continues on to $log\ n$n/2 butterfly operation groups, each of which operates on two sequence elements [29]. Because each group has separate processes, the FFT method may be computed. Every set of operations may be performed concurrently. The two items from the sequence are impacted at a time by each butterfly operation.

2.4. Self-Attention LSTM (SALSTM)

LSTM is a gated RNN design that uses certain gating mechanisms to control network information flow. This architecture is very useful for ECG time-series analysis since it reduces the vanishing and exploding gradient problems that are commonly seen in conventional RNN. By strengthening linear information pathways, the gating structure effectively stabilizes gradient propagation during training. The three control units, known as gates, present in LSTM, such as input gates, i_t, output gate$o_t$and forget gate f_t for AMI classification.

The i_t and f_t play a major role in LSTM's capacity to identify long-term dependencies.

Specifically, determining the level to which information from the input gate is allowed to current ECG input and network state is incorporated into the internal memory state, as defined in *equation (13),*

$$i_t = \sigma(U_i h_{t-1} + We_i x_t + b_i) \quad (13)$$

where *r* is the logistic function. At *t–1*, the memory block's output is denoted as h_{t-1}.

The input vector of the ECG at time *t* is denoted as x_t. We_i and U_i represent the weight matrix of the i_t,b_i constitutes the bias term of the i_t.The f_t uses *equation (14)* to decide the amount of previous information must be deleted,

$$f_t = \sigma\left(U_f h_{t-1} + We_f x_t + b_f\right) \quad (14)$$

The weight matrices for the f_t are represented as We_f and U_f. The bias term of the f_t is denoted as b_f. By using *equation (15)*, the amount of data that the internal state requires to communicate to the external state is determined by the output gate,

$$o_t = \sigma(U_o h_{t-1} + We_o x_t + b_o) \quad (15)$$

Where We_o and U_o are the weight matrices of the o_t, b_o is the bias term of the o_t.The gradient explosion issue is well resolved by LSTM. Over-fitting might result in the test set passing during the training phase.

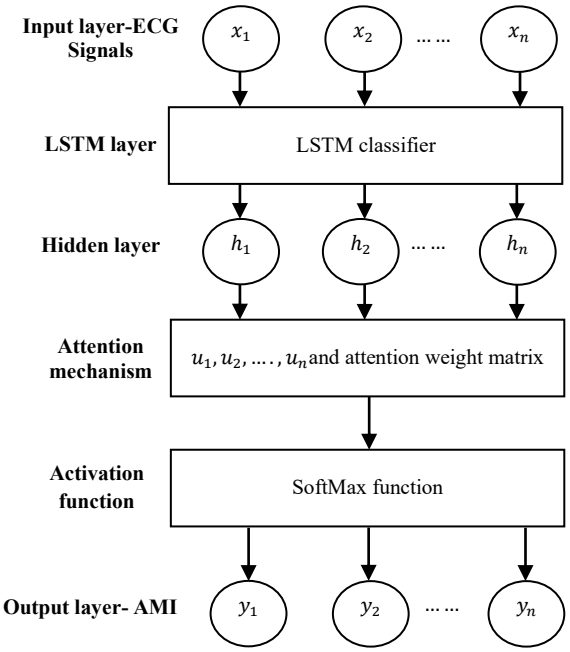

Figure 6. Structure of SALSTM Classifier

To address this issue, the LSTM's structure is optimized by adding a dropout layer. Multi-lead ECG input is high, standard LSTM may suffer from a dimension explosion problem, which reduces classification performance. To address this limitation, a self-attention mechanismis integrated to put emphasis on the most informative ECG leads and time segments. To implement the self-attention mechanism, the intermediate output of the LSTM encoder stores the input ECG signals.

Following that, a framework is trained for selectively learning these inputs, and when it generates an output, it connects the inputs and the output sequence [31-32]. *Figure 6* shows the framework of the SALSTM classifier.

The output $[h_1, h_2, h_3, \ldots h_n]$ nonlinear transformation of the LSTM produces $[u_1, u_2, u_3, \ldots, u_n]$. It will be the attention weight matrix generated by the SA mechanism $[a_1, a_2, a_3, \ldots, a_n]$, it may serve as a representation for each intermediate state's significance. The last step is to get the encoding vector *V* by performing input parameter and weight weighted sum. By decoding based on the *V*, the output *y* may be acquired. It is described by *equations (16-18)*,

$$u_k = tanh(We_k h_k + b_k) \quad (16)$$

$$a_k = \frac{exp\,(u_k^T u_s)}{\sum_{k=1}^{n} exp(u_k^T u_s)} \quad (17)$$

$$V = \sum_{k=1}^{n} a_k h_k \quad (18)$$

The weight matrix is denoted as b_k. Normalized attention weight is denoted as a_k. The attention matrix for a randomly initialised time series is denoted as u_s.Once the classification is done then detection is also confirmed with the clinical attributes.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

Algorithm 1: Proposed Framework
Input: *Noisy ECG signals and clinical parameters*
Output: *AMI classification result*
Step 1: ECG signals are collected from the PTB-XL dataset and clinical parameters from the clinical dataset.
Step 2: ECG signals are normalized by normalization and segmentation techniques.
Step 3: Noises such as Gaussian noise and baseline wander are introduced to generate noisy ECG signals.
Step 4: OVSAE model is initialized with parameters, including encoder–decoder layers, hidden neurons, latent vector dimensions, learning rate, and batch size.
Step 5: Layer-wise greedy pre-training is developed for stable parameter initialization.
Step 6: Encode noisy ECG signals into latent feature representations using the OVSAE encoder.
Step 7: Reconstruct noise removed ECG signals using the OVSAE decoder.
Step 8: Reconstruction loss is computed based on the MSE, MAE, and KL divergence.
Step 9: Model parameters are updated by the Adam optimizer until convergence criteria are satisfied.
Step 10: ECG features are extracted from the optimized OVSAE output.
Step 11: Denoised ECG sequences into the SALSTM classifier.
Step 12: Self-attention mechanism is introduced to recognize important temporal ECG features.
Step 13: Temporal dependencies are learned by LSTM hidden states.
Step 14: AMI and Healthy Control (HC) categorization is performed by Softmax activation.
Step 15: PSNR, MSE, MAE, and SSIM metrics has been computed for denoising model.
Step 16: SEN, SPE, ACC, and ERR metrics has been computed for classification.
Step 17: k-fold cross-validation is applied to confirm robustness and generalization capability.
Step 18: Generate final AMI prediction results.

3. RESULTS

The proposed model, and existing techniques were evaluated on a Windows 10 system with an i7-4790 CPU operating at 3.60GHz utilizing the Tensorflow and Keras framework in Python. The main experiment to evaluate the denoising techniques is described in this section, and classification methods for early MI prediction using Clinical Parameters in Risk Stratification, and PTB-XL ECG diagnostic dataset. OVSAE denoising method was compared to the various algorithms like DCNN-ED, FCN-DAE, and SDAE. OVSAE denoising method was compared to the various algorithms like DCNN-ED, FCN-DAE, and SDAE. SALSTM is compared with other methods like MCDANN, RTCN, MCA-net, and MLDA-RNN.

PTB-XL ECG Diagnostic Dataset: This is employed for ECG signal denoising and automated Acute Myocardial Infarction (AMI) detection. 21,837 records of 12-lead ECG signals from 18,885 individuals make up the extensive, publicly accessible clinical ECG dataset referred to as PTB-XL. High-resolution cardiac signals appropriate for DL-based analysis are produced by collecting each ECG record at 500 Hz for a length of 10 seconds.

3.1.1. Heart Attack Clinical Dataset

It was obtained from Kaggle. Various clinical parameters are included in this dataset for risk stratification. Patient-level clinical features associated with MI diagnosis, including clinical parameters in this dataset. The dataset provides binary output labels that indicate whether heart attack or not. Negative cases are encoded as 0 and positive cases as 1. To ensure consistency and facilitate model training, categorical and threshold-based features are normalized: male and female genders are encoded as 1 and 0, respectively, while the blood sugar attribute is binarized, with 120 mg/dL or more being recorded as 1 and 120 mg/dL or less being encoded as 0 in below *table 1*.

Table 1. Parameters and Experimental Settings of the Proposed Model

Category	Parameter	Description / Values
Dataset	ECG	PTB-XL ECG Diagnostic Dataset
	Number of ECG records	21,837
	Number of patients	18,885
	ECG leads	12-lead
	Signal duration	10 seconds
	Sampling frequency	500 Hz
	Clinical	Heart Attack Dataset (Kaggle)
	Number of attributes	clinical features + output
	Clinical parameters	Age, Gender, Heart Rate, Systolic BP, Diastolic BP, Blood Sugar, CK-MB, Troponin
	Output labels	Positive (1), Negative (0)
Noise Modeling	Noise types	Gaussian noise, baseline wander
	SNR levels	Variable (e.g., 0–20 dB)
Denoising Model	Model name	Optimized Variational Stacked Autoencoder (OVSAE)
	Encoder layers	Multiple nonlinear convolutional / dense layers
	Number of hidden layers	3-7 layers
	Hidden Neuron Size	64-512 neurons
	Latent vector dimension	16-128 features
	Latent space	Hierarchical latent representations
	Decoder	Symmetric to encoder
	Activation function	ReLU / Leaky ReLU
	Optimization strategy	Layer-wise greedy pre-training + global fine-tuning
Training	Optimizer	Adam

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

Parameters	Learning rate	0.001 (adjustable)
	Batch size	32 / 64
	Number of epochs	50-300 epochs
	Loss Convergence threshold	10^{-4}-10^{-6}
	Loss function	Reconstruction loss (MSE / MAE + KL divergence)
Classification Model	Classifier	Self-Attention-based LSTM (SALSTM)
	LSTM layers	1–2 layers
	Hidden units	128 / 256
	Self-attention	Temporal and multi-lead attention
	Dropout rate	0.3–0.5
	Activation function	Softmax
Search Strategy	Hyperparameter tuning	Iterative validation-based optimization
	Parameter selection	Classification accuracy
Computational Complexity	OVSAE training complexity	$O(n \times d \times h) - O(n \times h^2)$ Where n- Number of ECG samples used for training, d- Input feature dimension of ECG signals, h -Number of hidden neurons in encoder–decoder layers.
	SALSTM attention complexity	$O(T^2) - O(T \times h)$ Where T- Length of ECG temporal sequence / number of time steps.
Computational Cost	Memory utilization	4–12 GB
	Inference time per ECG sample	10–50 ms
Evaluation Metrics	Denoising metrics	PSNR, MSE, SSIM, MAE
	Classification metrics	Sensitivity, specificity, accuracy, and error

PSNR: PSNR evaluates the quality of the reconstructed ECG signal relative to the maximum possible signal power. It has been evaluated by *equation (19)* [33],

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right) \quad (19)$$

where MAX is denoted as the maximum possible amplitude of the ECG signal. Higher PSNR values correspond to improved denoising performance and signal fidelity.

MSE: Averaging the squared differences initial clean ECG data and the MSE is the measurement $x(i)$ and the denoised signal $\hat{x}(i)$. MSE is defined by *equation (20)* [33],

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(x(i) - \hat{x}(i)\right)^2 \quad (20)$$

The accurate reconstruction and less noise distortion are indicated by lower MSE values.

MAE: It determines the average of the actual x,and denoised ECG signals $\hat{x}$ by *equation (21)* [33],

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|x(i) - \hat{x}(i)| \quad (21)$$

SSIM: It evaluates the perceptual similarity between the actual x, and denoised ECG signals $\hat{x}$ using *equation (22)* to consider structural information, contrast, and brightness [33],

$$SSIM(x,\hat{x}) = \frac{(2\mu_x\mu_{\hat{x}}+C_1)(2\sigma_x\sigma_{\hat{x}}+C_2)}{(\mu_x^2+\mu_{\hat{x}}^2+C_1)(\sigma_x^2+\sigma_{\hat{x}}^2+C_2)} \quad (22)$$

where $\mu_x, \mu_{\hat{x}}$ is denoted as the mean of actual x, and denoised ECG signals $\hat{x}$,$\sigma_x, \sigma_{\hat{x}}$ is denoted as the standard deviation of actual x, and denoised ECG signals $\hat{x}$, C_1, C_2 is denoted as the constants.

SEN, SPE, ACC, and ERR metrics are calculated based on the confusion matrix like True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) [34].

Sensitivity (SEN): Its sensitivity indicates how efficiently the model detects positive cases (AMI present). It is defined by *equation (23),*

$$\text{SEN} = \frac{\text{TP}}{\text{TP+FN}} \times 100 \quad (23)$$

Specificity (SPE): The accurateness of the framework assesses its ability to accurately detect negative instances (AMI absent). It is defined by *equation (24)*,

$$\text{SPE} = \frac{\text{TP}}{\text{TP+FP}} \times 100 \quad (24)$$

Accuracy (ACC): ACC measures model accuracy across positive and negative classes. It is defined by *equation (25),*

$$\text{ACC} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \times 100 \quad (25)$$

Error (ERR): Error represents the proportion of incorrect predictions that have been computed by *equation (26),*

$$\text{ERR} = \frac{\text{FP+FN}}{\text{TP+TN+FP+FN}} \times 100 \quad (26)$$

The PTB database was subjected to five-fold cross-validation. Following that, *table 2* shows the confusion matrix across the folds.

Table 2. Confusion Matrix of PTB Database

True label/ predicted label	AMI	HC
AMI	25182	1982
HC	1565	23691

Table 3 shows the confusion matrix comparison of all classifiers such as MCDANN, RTCN, MCA-Net, MLDA-RNN, and SALSTM with total number of samples as 52420. From records like 25182, 23691, 1987, and 1565, proposed classifier has accurately detected as 25176, 23685, 1987, and 1572 records.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

Table 3. Confusion Matrix of Classifiers

METHODS	TP	TN	FP	FN	TOTAL
MCDANN	23535	22018	3589	3278	52420
RTCN	23924	22627	2914	2955	52420
MCA-net	24596	23021	2573	2230	52420
MLDA-RNN	24952	23254	2280	1934	52420
SALSTM	25176	23685	1987	1572	52420

Table 4 shows the denoising methods results of DCNN-ED, FCN-DAE, SDAE, and OVSAE with respect to PSNR, MSE, MAE, and SSIM. The proposed classifier has highest PSNR, and SSIM of 45.79 dB, and 0.9283. It also has lowest MSE, and MAE of 0.0053, and 0.049.

Table 4. Denoising Comparison of Classification Methods

Methods	PSNR (dB)	MSE	MAE	SSIM
DCNN-ED	36.43	0.0145	0.096	0.8542
FCN-DAE	38.92	0.0118	0.083	0.8795
SDAE	42.66	0.0079	0.065	0.9074
OVSAE	45.79	0.0053	0.049	0.9283

Table 5. Comparison of Classification Methods

Methods	SEN (%)	SPE (%)	ACC (%)	ERR (%)
MCDANN	87.78	85.98	86.90	13.09
RTCN	89.01	88.60	88.81	11.20
MCA-net	91.69	89.95	90.84	9.16
MLDA-RNN	92.81	91.07	91.96	8.04
SALSTM	94.13	92.26	93.21	6.79

Table 5 shows the classification methods results of MCDANN, RTCN, MCA-Net, MLDA-RNN, and SALSTM with respect to SEN, SPE, ACC, and ERR. The proposed classifier has highest SEN, SPEC, and ACC of 94.13%, 92.26%, and 93.21%. It also has lowest ERR of 6.79%.

The k-fold cross-validation results demonstrate that the proposed SALSTM model maintains consistent classification performance across different dataset partitions. The low variation among folds indicates strong generalization capability and robustness of the proposed framework. Table 6 obtained average accuracy of 93.24% and low error rate of 6.76% confirm the effectiveness of the OVSAE-SALSTM model for reliable AMI detection under varying ECG signal conditions.

Table 6. K-Fold Cross-Validation Results of SALSTM Model

Fold	SEN (%)	SPE (%)	ACC (%)	ERR (%)
Fold 1	93.82	91.94	92.96	7.04
Fold 2	94.07	92.11	93.15	6.85
Fold 3	94.31	92.45	93.48	6.52
Fold 4	94.22	92.37	93.39	6.61
Fold 5	94.13	92.26	93.21	6.79
Average	94.11	92.23	93.24	6.76

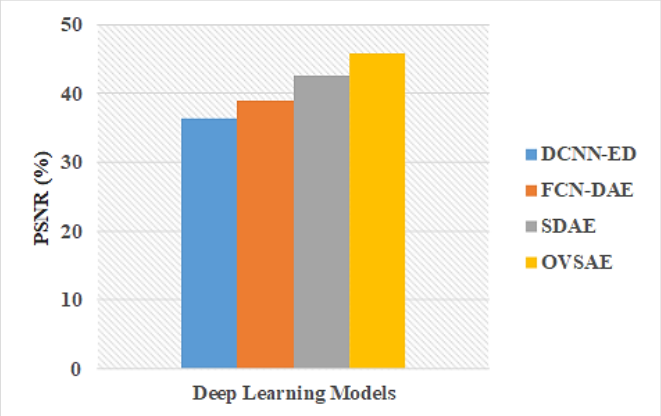

Figure 7. PSNR Comparison of Denoising Models

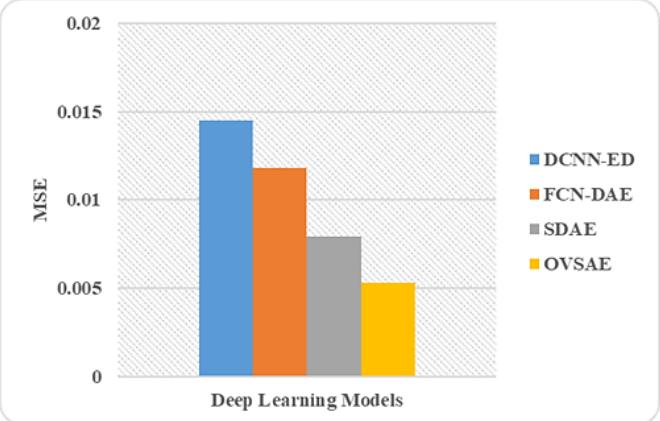

Figure 8. MSE Comparison of Denoising Models

Figure 7 illustrates the PSNR performance of four reconstruction methods like DCNN-ED, FCN-DAE, SDAE, and OVSAE. DCNN-ED achieves the lowest PSNR value of 36.43 dB, indicating relatively lower reconstruction quality. FCN-DAE shows an improvement with a PSNR of 38.92 dB, reflecting enhanced noise reduction and signal preservation. A substantial performance improvement is observed with SDAE, which attains a PSNR of 42.66 dB, demonstrating superior reconstruction ACC due to its better learning capability. At 45.79 dB, OVSAE attains the maximum PSNR value. This score shows a strong signal with little distortion. In comparison to DCNN-ED, FCN-DAE, and SDAE, the suggested strategy achieves high 9.36 dB, 6.87 dB, and 3.13 dB. Unlike conventional denoising models, the proposed OVSAE employs variational latent feature learning combined with layer-wise greedy pre-training and global fine-tuning to improve ECG signal reconstruction quality and robustness against noise.

Figure 8 displays the MSE values for DCNN-ED, FCN-DAE, SDAE, and OVSAE. The highest MSE of 0.0145 is recorded by DCNN-ED. With the MSE of 0.0118, FCN-DAE demonstrates a decrease in ERR, indicating better reconstruction ACC. This high MSE indicates a greater signal reconstruction ERR. With an MSE of 0.0079, SDAE exhibits a

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

significant reduction in MSE, indicating improved feature representation and efficient noise suppression. With the lowest MSE value of 0.0053, OVSAE demonstrates excellent ECG signal quality and low reconstruction error. In comparison to DCNN-ED, FCN-DAE, and SDAE, the suggested system has the lowest MSE of 0.0092, 0.0065, and 0.0026.

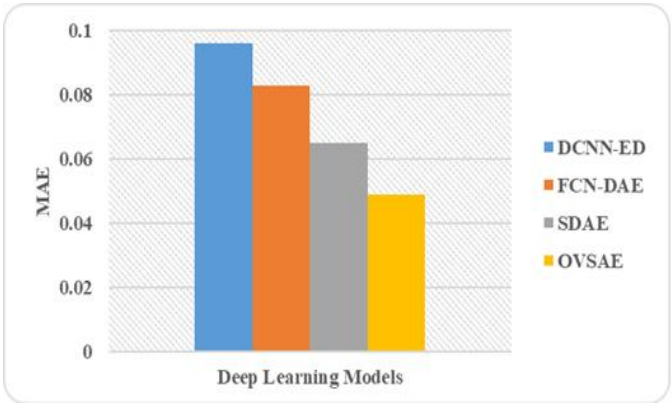

Figure 9. MAE Comparison of Denoising Models

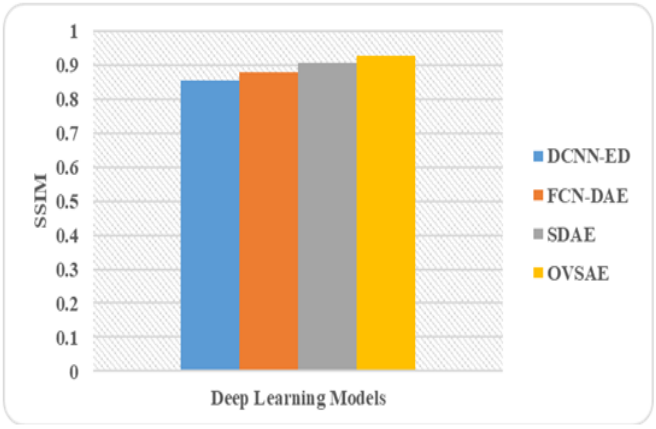

Figure 10. SSIM Comparison of Denoising Models

Figure 9 illustrates the comparison of MAE values for four DL models like DCNN-ED, FCN-DAE, SDAE, and OVSAE. DCNN-ED has the highest MAE value (0.096), FCN-DAE shows moderate improvement with a MAE of 0.083. SDAE further reduces the error to 0.065, demonstrating better learning capability. Among all the models, OVSAE achieves the lowest MAE value of 0.049, highlighting its superior performance and effectiveness in minimizing reconstruction error. The suggested method has 0.047, 0.034, and 0.016 lowest MAE when compared to DCNN-ED, FCN-DAE, and SDAE.

SSIM comparisons for 4DL models are DCNN-ED, FCN-DAE, SDAE, and OVSAE is illustrated in figure 10. DCNN-ED achieves an SSIM of 0.8542, FCN-DAE shows enhanced efficiency using an SSIM value of 0.8795. SDAE further enhances the structural similarity, reaching an SSIM of 0.9074. Notably, OVSAE attains the highest SSIM value of 0.9283, demonstrating its superior capability in preserving structural information and producing high-quality reconstructed outputs. The suggested method has 0.0741, 0.0488, and 0.0209 highest SSIM when compared to DCNN-ED, FCN-DAE, and SDAE.

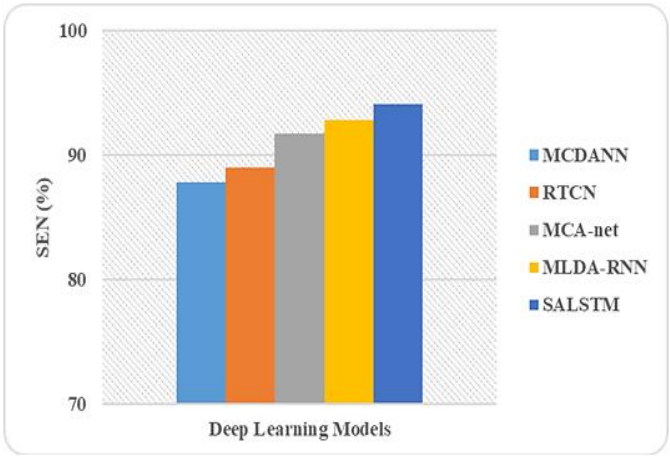

Figure 11. SEN Comparison of Deep Learning Models

DL model comparison for SEN (MCDANN, RTCNA, MCA-net, MLDA-RNN, and SALSTM) is illustrated in figure 11. The suggested classifier has the highest SEN outcomes of 94.13%, the other classifiers like RTCN, MCA-net, MLDA-RNN, and SALSTM gives lowest SEN results of 87.78%, 89.01%, 91.69%, and 92.81%. The suggested method has 6.35%, 5.12%, 2.44%, and 1.32% higher SEN when compared to RTCN, MCA-net, MLDA-RNN, and SALSTM (*Refer Table 5*).

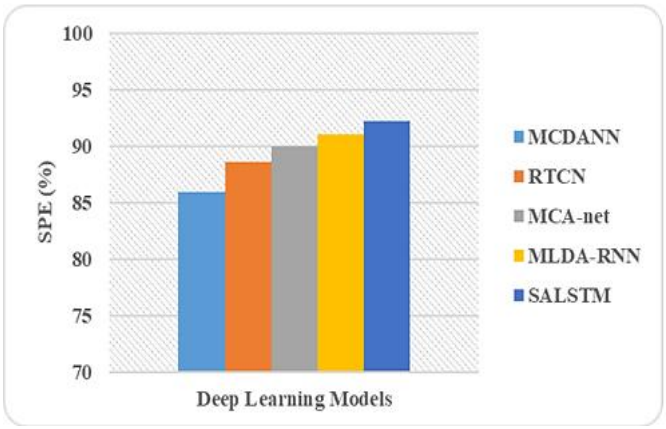

Figure 12. SPE Comparison of Deep Learning Models

DL model comparison using SPE (MCDANN, RTCN, MCA-net, MLDA-RNN, and SALSTM) is illustrated in figure 12. The suggested classifier has the highest SPE results of 92.26%, the other classifiers like RTCN, MCA-net, MLDA-RNN, and SALSTM gives lowest SPE results of 85.98%, 88.60%, 89.95%, and 91.07%. The suggested system has 6.28%, 3.66%, 2.31%, and 1.19% higher SPE when compared to RTCN, MCA-net, MLDA-RNN, and SALSTM (*Refer Table 5*).

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 14, Issue 2 | Pages 361-374| e-ISSN: 2347-470X

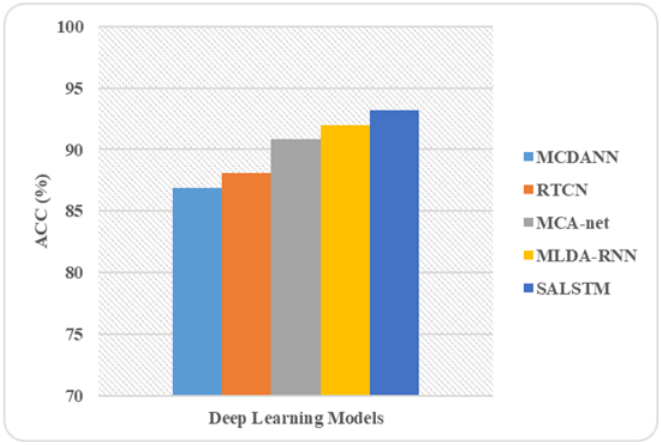

Figure 13. ACC Comparison of Deep Learning Models

Comparing DL models using ACC (MCDANN, RTCN, MCA-net, MLDA-RNN, and SALSTM) is illustrated in figure 13. The suggested classifier has the highest ACC results of 93.21%, the other classifiers like RTCN, MCA-net, MLDA-RNN, and SALSTM gives lowest ACC results of 86.90%, 88.11%, 90.84%, and 91.96%. The suggested system has 6.31%, 5.10%, 2.37%, and 1.25% higher ACC when compared to RTCN, MCA-net, MLDA-RNN, and SALSTM (*Refer Table 5*).

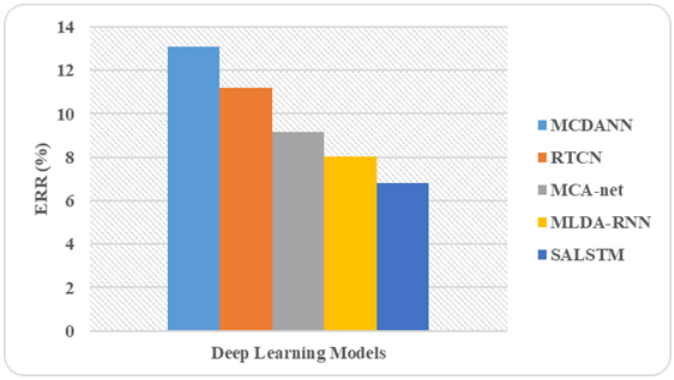

Figure 14. ERR Comparison of Deep Learning Models

DL model ERR comparison (MCDANN, RTCN, MCA-net, MLDA-RNN, and SALSTM) is illustrated in figure 14. The suggested classifier has the lowest ERR results of 6.79%, the other classifiers like RTCN, MCA-net, MLDA-RNN, and SALSTM gives increased ERR results of 13.09%, 11.20%, 9.16%, and 8.04%. The suggested system has 6.31%, 5.10%, 2.37%, and 1.25% lower ERR when compared to RTCN, MCA-net, MLDA-RNN, and SALSTM (Refer Table 5). In addition, the proposed SALSTM classifier effectively captures temporal dependencies and important ECG patterns using the self-attention mechanism, resulting in enhanced AMI classification performance.

Discussion and findings: The PSNR value of 45.79 dB and SSIM value of 0.9283 obtained by OVSAE were better than the other methods, indicating the better signal reconstruction quality and structural preservation. Furthermore, the model produced the lowest MSE (0.0053) and MAE (0.049) implying efficient noise suppression and reduced reconstruction error. These results verify that the optimized variational stacked autoencoder is able to learn robust latent representations and preserve important ECG morphological characteristics even under noisy conditions.

The proposed method achieved the highest Sensitivity (94.13%), Specificity (92.26%) and Accuracy (93.21%) and the lowest Error Rate (6.79%). The results demonstrate that the combination of the self-attention mechanism and LSTM is effective in capturing significant temporal dependencies and discriminative ECG features related to AMI. The model exhibited a good performance on all the folds with an average Accuracy of 93.24% and Error Rate of 6.76% that indicated the stability of the model over different dataset splits. The low variance across folds indicates the reliability of the proposed OVSAE-SALSTM framework and less susceptibility to overfitting.

4. CONCLUSION

In this paper, leveraging multimodal datasets that include ECG signals and clinical parameters dataset can improve diagnostic accuracy and robustness. The OVSAE model combines the generative capability of Variational Autoencoders (VAEs) with the hierarchical learning strength Stacked Autoencoders (SAEs) to efficiently remove ECG noise and preserve important cardiac features. Self-Attention Long Short-Term Memory (SALSTM) classifier effectively captures latent representations and temporal dependencies in ECG signals through the self-attention mechanism, thereby improving AMI detection performance. The LSTM encoder's intermediate output retains the input ECG signals. Thus, by combining the output sequences for AMI and HC, the method gets trained to selectively learn these inputs. Experimental results demonstrated superior denoising performance with PSNR, MSE, MAE, and SSIM values of 45.79 dB, 0.0053, 0.049, and 0.9283, respectively. AMI results have been measured using metrics like SEN, SPE, ACC, and ERR as 94.13%, 92.26%, 93.21%, and 6.79% for proposed classifier. The Future work will focus on integrating explainable AI techniques such as Saliency Maps and Score-CAM (Score-CAM) to improve clinician interpretability and physician trust of the ECG signal such as the ST-segment and T-waves. In addition, future studies will investigate improved NSTEMI detection using multimodal clinical data and validate the proposed framework through prospective real-world clinical studies. Moving from retrospective trials to real-world, prospective studies to verify the clinical impact on patient results.

Funding Source: This research received no external funding.

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

REFERENCES

[1] Siontis, K.C.; Noseworthy, P.A.; Attia, Z.I.; Friedman, P.A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nature Reviews Cardiology* 2021, *18*, 465–478.

[2] Moras, E.; Yakkali, S.; Gandhi, K.D.; Virk, H.U.H.; Alam, M.; Zaid, S.; Barman, N.; Jneid, H.; Vallabhajosyula, S.; Sharma, S.K; Krittanawong, C. 2024. Complications in acute myocardial infarction: Navigating challenges in diagnosis and management. *Hearts* 2024, *5(1)*, 122-141.

[3] Scheldeman, L.; Sinnaeve, P.; Albers, G.W.; Lemmens, R.; Van de Werf, F. Acute myocardial infarction and ischaemic stroke: differences and similarities in reperfusion therapies—a review. *European Heart Journal* 2024, *45(30)*, 2735-2747.

[4] Lindahl, B.; Mills, N.L.2023. A new clinical classification of acute myocardial infarction. *Nature Medicine* 2023, *29(9)*, 2200-2205.

[5] Siontis K. C.; Noseworthy P. A.; Attia Z. I.; Friedman P. A. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nature Reviews Cardiology* 2021,*18*, 465–478.

[6] Lee, Y.; Ahn, J.H.; Yang, H.M. Artificial intelligence-enhanced electrocardiography for acute myocardial infarction detection: a systematic review. *Cardiovascular Diagnosis and Therapy 2026*, 16(2), pp.1-24.

[7] Gupta, V.; Hilgendorf, L.; Andersson, E.; Louca, A.; Shahmari, A.; Hjalmarsson, A.; Saini, R.; Pirazzi, C.; Alchay, M.; Rawshani, A.Multimodal deep learning for acute myocardial infarction detection from 12-lead electrocardiogram: a multi-centre study with cross-hospital validation. *European Heart Journal - Digital Health* 2026, *7(2)*, 1-11.

[8] Bhatt, D.L.; Lopes, R.D.; Harrington, R.A. Diagnosis and treatment of acute coronary syndromes: a review. *Jama* 2022, *327(7)*, 662-675.

[9] Al-Zaiti, S.S.; Martin-Gill, C.; Zègre-Hemsey, J.K.; Bouzid, Z.; Faramand, Z.; Alrawashdeh, M.O.; Gregg, R.E.; Helman, S.; Riek, N.T.; Kraevsky-Phillips, K.; Clermont, G. Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. *Nature Medicine* 2023, *29(7)*, 1804-1813.

[10] Ali, S.N.; Shuvo, S.B.; Al-Manzo, M.I.S.; Hasan, A.; Hasan, T. An end-to-end deep learning framework for real-time denoising of heart sounds for cardiac disease detection in unseen noise. *IEEE Access* 2023, *11*, 87887-87901.

[11] Rasti-Meymandi, A.; Ghaffari, A. A deep learning-based framework For ECG signal denoising based on stacked cardiac cycle tensor. *Biomedical Signal Processing and Control* 2022, *71*,103275.

[12] Qiang, Y.; Dong, X.; Yang, Y.Automatic detection and localisation of myocardial infarction using multi-channel dense attention neural network. *Biomedical Signal Processing and Control* 2024, *89*,105766.

[13] Gunawan, G.; Akbar, A.A.;Andriani, W., Application of deep neural network with stacked denoising autoencoder for ECG signal classification. *Journal of Intelligent Decision Support System (IDSS)* 2024, *7(2)*, 173-187.

[14] Chiang, H.T.; Hsieh, Y.Y.; Fu, S.W.; Hung, K.H. Tsao, Y.; Chien, S.Y. Noise reduction in ECG signals using fully convolutional denoising autoencoders. *IEEE Access* 2019, *7*, 60806-60813.

[15] Fotiadou, E.; Vullings, R.Multi-channel fetal ECG denoising with deep convolutional neural networks. *Frontiers in Pediatrics* 2020, *8*, 1-13.

[16] Prabhakararao, E.; Dandapat, S., 2020. Myocardial infarction severity stages classification from ECG signals using attentional recurrent neural network. *IEEE Sensors Journal* 2020, *20(15)*, 8711-8720.

[17] Jahmunah, V.; Ng, E.Y.; Tan, R.S.; Oh, S.L.; Acharya, U.R.Explainable detection of myocardial infarction using deep learning models with Grad-CAM technique on ECG signals. *Computers in Biology and Medicine* 2022, *146*, 1-46.

[18] Chen, Y.; Gao, Q.; Ye, J.; Li, Y.; Wan, X.Data Augmentation-Enhanced Myocardial Infarction Classification and Localization Using a ResNet-Transformer Cascaded Network. *Biology* 2025, *14(10)*,1-25.

[19] Mahendran, R.K.; Prabhu, V.; Parthasarathy, V.; Mary Judith, A. Deep learning based adaptive recurrent neural network for detection of myocardial infarction. *Journal of Medical Imaging and Health Informatics* 2021, *11(12)*, 3044-3053.

[20] Nurmaini, S.; Tondas, A.E.; Darma Wahyuni, A.; Rahmatullah, M.N.; Partan, R.U.; Firdaus, F.; Tutuko, B.; Pratiwi, F.; Juliano, A.H.; Khoirani, R. Robust detection of atrial fibrillation from short-term electrocardiogram using convolutional neural networks. *Future Generation Computer Systems* 2020, *113*, 304-317.

[21] Mirza, A.H.; Nurmaini, S.; Partan, R.U. Automatic classification of 15 leads ECG signal of myocardial infarction using one-dimension convolutional neural network. *Applied Sciences* 2022, *12(11)*,1-13.

[22] Khatar, Z.; Bentaleb, D.; El Mansouri, M. Integrating Advanced Combined Numerical Filters for ECG Denoising and Cardiovascular Disease Classification Using Deep Learning. *In International Conference on Digital Technologies and Applications* 2024, 539-547.

[23] Pan, W.; An, Y.; Guan, Y.; Wang, J. MCA-net: A multi-task channel attention network for Myocardial infarction detection and location using 12-lead ECGs. *Computers in Biology and Medicine* 2022, *150*, 106199.

[24] Zhao, T.; Cui, Y.; Ji, T.; Luo, J.; Li, W.; Jiang, J.; Gao, Z.; Hu, W.; Yan, Y.; Jiang, Y.; Hong, B.; VAEEG: Variational auto-encoder for extracting EEG representation. *NeuroImage* 2024, *304*, 1-11.

[25] Wang, Y.; Qiu, S.; Li, D.; Du, C.; Lu, B.L.; He, H.; Multi-odal domain adaptation variational autoencoder for EEG-based emotion recognition. *IEEE/CAA Journal of Automatica Sinica* 2022, *9(9)*, 1612-1626.

[26] Lv, J.; Yang, H.; Li, P. Wasserstein distance rivals kullback-leibler divergence for knowledge distillation. *Advances in Neural Information Processing Systems* 2024, *37*, 65445-65475.

[27] Bonnici, V.A. Maximum value for the Kullback–Leibler divergence between quantized distributions. *Information* 2024, *15(9)*, 1-22.

[28] Leitersdorf, O.; Boneh, Y.; Gazit, G.; Ronen, R.; Kvatinsky, S.FourierPIM: High-throughput in-memory Fast Fourier Transform and polynomial multiplication. *Memories-Materials, Devices, Circuits and Systems* 2023, *4*,1-8.

[29] Chatterjee, K.; Kumar, S.S.; Kumar, R.P.; Bandyopadhyay, A.; Swain, S.; Mallik, S.; Al-Rasheed, A.; Abbas, M.;Soufiene, B.O. Future air quality prediction using long short-term memory based on hyper heuristic multi-chain model. *IEEE Access* 2024, *12*, 123678 – 123693.

[30] Waheed, W.; Xu, Q.; Aurangzeb, M.; Iqbal, S.; Dar, S.H.; Elbarbary, Z.M.S. Empowering data-driven load forecasting by leveraging long short-term memory recurrent neural networks. *Heliyon* 2024, *10(24)*, 1-15.

[31] Luo, T.; Cao, X. D.; Li, J.; Dong, K.; Zhang, R.; Wei, X. L. Multi-task prediction model based on ConvLSTM and encoderdecoder. *Intelligent Data Analysis* 2021, *25(2)*, 359–382.

[32] Shi, M. J.; Yang, B. H.; Chen, R.; Ye, D. S.Logging curve prediction method based on CNN-LSTM-attention. *Earth Science Informatics* 2022, *15*, 2119–2131.

[33] Mao, J.; Li, H.; Zhao, Y.An innovative deep learning-driven technique for restoration of lost high-density surface electromyography signals. *Applied Intelligence* 2025, *55(7)*,1-16.

[34] Rainio, O.; Teuho, J.; Klén, R.Evaluation metrics and statistical tests for machine learning. *Scientific Reports* 2024, *14(1)*,1-14.