MED-Net: Multi-Scale Enhanced Dual Temporal Convolution Attention Network

G700

A

PROJECT SCHOOL

REPORT ON

MED-Net: Multi-Scale Enhanced Dual Temporal Convolution Attention Network

Submitted By

Ehsaas Nahata	245523748018
Honeysha Magam	245523748308
Kandibanda Sathwika	245523733026
Kanishka Gupta	245523733027
Madhuri Vemparala	245523753051
Neha Thapasvi Kodithala	245523753025

Under the guidance

of

KVN Hari Babu

Assistant Professor

KESHAV MEMORIAL ENGINEERING COLLEGE

Kachavanisingaram Village, Hyderabad, Telangana 500058.

December, 2025

KESHAV MEMORIAL ENGINEERING COLLEGE
A Unit of Keshav Memorial Technical Education (KMTES)
Approved by AICTE, New Delhi & Affiliated to Osmania University, Hyderabad

CERTIFICATE

This is to certify that the project work entitled **"MED-Net: Multi-Scale Enhanced Dual Temporal Convolution Attention Network"** *is a bonafide work carried out by the following students of IIInd Year V Semester* **Bachelor of Engineering in CSE/CSE(AIML)/CS** *during the academic year 2025–2026 and is a record of bonafide work carried out by them.*

Ehsaas Nahata	245523748018
Honeysha Magam	245523748308
Kandibanda Sathwika	245523733026
Kanishka Gupta	245523733027
Madhuri Vemparala	245523753051
Neha Thapasvi Kodithala	245523753025

Project Mentor

KVN Hari Babu

Assistant Professor

Contents

Abstract

Electrocardiographic (ECG) signals are essential for evaluating cardiac activity and detecting abnormalities such as arrhythmias and atrial fibrillation. Automated ECG analysis systems are particularly important in long-term monitoring scenarios, including 24-hour Holter recordings, where manual interpretation is inefficient and error-prone. This project presents MED-Net, an ECG anomaly detection and diagnosis framework based on the Dual Temporal Convolutional Network Attention for Anomaly Detection (DTAAD) architecture, designed to accurately analyze multivariate ECG time-series data while remaining computationally efficient for real-world clinical use.

MED-Net integrates an autoregressive (AR) prediction model with an autoencoder (AE) to capture both temporal dependencies and reconstruction errors in ECG signals. To further enhance anomaly discrimination, scaling strategies and feedback mechanisms are introduced to improve prediction accuracy and amplify correlation differences between normal and abnormal patterns. The core Dual TCN-Attention (DTA) network employs dual Transformer-based attention layers, enabling precise temporal feature extraction within an ultra-lightweight model. Experimental results on public ECG datasets demonstrate that MED-Net achieves superior detection and diagnostic performance, improving F1-scores by approximately 1.5% over DTAAD and 5% over convolutional autoencoders, while reducing training time by 99% compared to CAE and 50% compared to DTAAD.

Beyond model development, MED-Net is implemented as a full-stack medical AI application to support end-to-end ECG analysis. The system features a React-based frontend for interactive visualization, a Node.js and Express backend for authentication and data management, and a FastAPI-based Python model server built with PyTorch for real-time inference. A PostgreSQL database via prisma stores user data, predictions, and chat logs, while Groq SDK is integrated to provide clinical interpretations and decision support. This scalable architecture enables efficient deployment of MED-Net as an intelligent ECG monitoring and clinical decision support system.

List of Figures

List of Tables

Chapter 1

Introduction

Anomaly detection refers to the identification of observations that deviate significantly from expected data patterns and are therefore labeled as anomalies. Multivariate time-series data are ubiquitous in real-world systems, making anomaly detection in such settings a long-standing research focus [8]. These techniques are widely applied across diverse domains including finance, healthcare, cyber security, industrial monitoring, power systems, transportation, robotics, urban IoT, and sensor networks. Consequently, anomaly detection has become a prominent topic in data mining, probabilistic statistics [2], machine learning [5], and computer vision [15].

In recent years, Deep Learning (DL) based anomaly detection methods have achieved remarkable progress. DL models excel at learning complex representations from temporal, streaming, spatial, and high-dimensional data, overcoming limitations of traditional learning approaches. Deep anomaly detection leverages deep neural networks to learn discriminative features or anomaly scores, typically within prediction-based frameworks [21]. These methods have demonstrated strong performance in modeling non-linear temporal dependencies, making them particularly suitable for time-series anomaly detection tasks.

Electrocardiogram (ECG) signals play a crucial role in assessing cardiac activity and heart health, especially for identifying conditions such as arrhythmias and atrial fibrillation. ECG signals record the electrical activity of the heart over durations ranging from short clinical recordings (10 seconds) to long-term continuous monitoring using Holter devices. Heart activity is measured across multiple channels, typically ranging from 1 to 12 leads, with some devices supporting vectorcardiograms (VCG) to capture both magnitude and direction of cardiac electrical forces. Recent advancements in wearable technology have further enabled real-time ECG monitoring during daily activities.

Despite their clinical importance, ECG signals present significant challenges for automated analysis. Large-scale time-series databases generated from medical devices often

exhibit uncertainty, temporal variability, and noise. Manual ECG interpretation by clinicians is time-consuming and subjective, particularly for long-term recordings. Although experts can visually identify rhythm irregularities, such as arrhythmias, computer-aided diagnostic tools are essential for scalable and consistent anomaly detection. Consequently, automated ECG anomaly detection has emerged as a critical research direction to assist clinicians with minimal human intervention.

Modern anomaly detection systems must also address challenges arising from massive data volume, high dimensionality, real-time inference requirements, and scarcity of labeled data. The proliferation of IoT devices has increased both data diversity and complexity, complicating reliable analysis. While supervised models are limited by annotation costs, unsupervised approaches offer practical alternatives [3]. Additionally, beyond merely detecting anomalies, identifying their root causes is increasingly important, requiring multi-category predictions that capture both anomaly presence and origin. Addressing these challenges remains central to advancing multivariate time-series anomaly detection in real-world healthcare applications.

Chapter 2

Literature Survey

Anomaly detection in multivariate time-series data, particularly ECG signals, remains a challenging research problem due to high dimensionality, strong temporal dependencies, limited labeled data, and real-time inference requirements. Traditional unsupervised techniques based on statistical modeling, distance metrics, or density estimation often fail to capture the complex temporal dynamics inherent in physiological signals. While deep learning methods have significantly advanced ECG anomaly detection, many existing architectures suffer from inefficiencies in modeling long-range dependencies, computational overhead, and unstable training behavior.

A large body of work relies on recurrent architectures such as RNNs and LSTMs to model temporal correlations [22, 10]. Cloud-based LSTM Autoencoder frameworks combined with wavelet scattering [18] reduce signal dimensionality but introduce sequential bottlenecks and dependency on external infrastructure. Hybrid architectures integrating BiGRU–BiLSTM, dilated CNNs, and GAN-based augmentation [20] improve accuracy at the cost of increased model complexity, training time, and overfitting risk. Similarly, GRU-based encoder–decoder models such as OmniAnomaly [26] struggle with long ECG sequences and abrupt noise variations, while rigid segmentation-based approaches [6] lack adaptability to variable-length signals.

To overcome recurrent inefficiencies, CNN- and Transformer-based approaches have been explored [28]. Convolutional autoencoders and 1D CNN models [24, 14] offer faster inference but are often limited to local temporal contexts and fixed signal configurations. QRS-focused CNN methods [29] fail to capture global rhythm patterns, which are crucial for arrhythmia detection. Transformer-based anomaly detection models such as TranAD [27] and masked reconstruction frameworks [4] improve parallelism but incur high computational costs and often emphasize short temporal windows, missing long-range dependencies.

Generative and adversarial approaches attempt to model normal ECG distributions but suffer from training instability and high resource demands. Adversarial autoencoder frameworks combining TCNs with GANs [23] require careful loss balancing, while memory-augmented GAN-based systems such as MadeGAN [13] introduce significant computational and memory overhead. Other methods relying on fuzzy clustering or manual visual inspection [12, 19, 30] lack robustness and scalability in noisy, real-time clinical environments.

To address these limitations, the proposed MED-Net framework introduces a lightweight, non-adversarial architecture built upon Dual Temporal Convolutional Networks [1, 31]. By integrating efficient multi-scale temporal feature extraction, lightweight attention mechanisms, and fixed fusion strategies, MED-Net eliminates recurrent bottlenecks, captures both local and global temporal dependencies, and ensures stable training. Unlike rigid segmentation or GAN-based approaches, MED-Net incorporates dynamic decoders and flexible label handling, making it adaptable to varying ECG lengths and preprocessing strategies, while achieving improved speed, stability, and diagnostic precision.

Chapter 3

Proposed Work, Architecture & Implementation Details

3.1. Proposed Work

In this chapter we present MED-Net, a morphology-enhanced deep network for ECG anomaly detection that extends the Dual-TCN and attention-based DTAAD architecture [31]. The core idea is to preserve the autoregressive modeling capability of DTAAD while tailoring the feature extraction and attention mechanisms to ECG morphology. MED-Net incorporates a dual temporal convolutional backbone, a multi-scale temporal feature extractor designed to capture P, QRS and T waves, a lightweight ECG-specific attention block, an enhanced attention refinement module, a fixed-weight fusion scheme for combining multiple feature streams, an increased temporal window to capture longer cardiac cycles, and a learning-rate decay schedule to improve convergence stability. The overall system is trained in a self-supervised reconstruction framework, so that deviations from normal ECG dynamics manifest as increased reconstruction error and can be used for both anomaly detection and diagnosis.

3.2. Problem Definition

The ECG signal is modeled as a univariate or multivariate time series. Let

$$T = \{x_1, x_2, \ldots, x_T\}, \qquad x_t \in R^C, \tag{1}$$

where T denotes the total number of time steps and C is the number of channels or leads. Under an autoregressive assumption, the joint distribution of the ECG sequence can be factorized as

$$p(x_{1:T}) = \prod_{t=1}^{T} p(x_t \mid x_{1:t-1}), \tag{2}$$

so that each sample depends on its preceding history. The goal of MED-Net is to learn a model that accurately captures the normal temporal dynamics and morphology of this process. Once trained, large deviations in prediction or reconstruction with respect to the learned normal pattern are interpreted as anomalies. Formally, given an observed prefix $x_{1:T}$, the model either reconstructs the same window or forecasts future values, and the discrepancy between the model output and the true signal is used to compute anomaly scores over time.

3.3. Data Preprocessing and Windowing

The raw ECG recordings may contain noise, missing values and artefacts caused by sensor drift or motion. To obtain a clean input representation, we first retain only numeric channels corresponding to ECG amplitudes and discard non-informative metadata such as timestamps or identifiers. Invalid or non-numeric entries are coerced to `NaN`, rows or columns that are entirely empty are removed, and remaining missing samples are imputed, for example by zero filling or interpolation. After cleaning, each recording can be represented as a matrix $X \in R^{T \times C}$, where each row corresponds to a time step and each column to a channel.

To stabilize training and ensure all channels are on a comparable scale, the data are normalized using feature-wise statistics computed on the training portion. For channel c, with mean μ_c and standard deviation σ_c, z-score normalization is applied as

$$\tilde{x}_{t,c} = \frac{x_{t,c} - \mu_c}{\sigma_c}, \qquad t = 1, \ldots, T, \;\; c = 1, \ldots, C. \tag{3}$$

This transformation reduces the effect of varying signal amplitudes across subjects and leads and improves numerical stability of deep optimization.

After normalization, the continuous ECG is partitioned into fixed-length windows that form the basic training and inference units. For a window length L, the i-th window is defined as

$$W_i = [\tilde{x}_{i,:}, \; \tilde{x}_{i+1,:}, \; \ldots, \; \tilde{x}_{i+L-1,:}] \in R^{L \times C}, \tag{4}$$

where $\tilde{x}_{t,:}$ denotes the length-C vector at time t. In practice, overlapping windows are

12

employed to preserve temporal continuity and enable fine-grained localization of anomalous segments. For compatibility with convolutional and attention layers, each window is then rearranged into channel-first format, $W_i \in R^{C \times L}$, so that channels index rows and time indexes columns.

3.4. MED-Net Architecture

The MED-Net architecture is designed to capture both local ECG morphology and long-range temporal dependencies. An overview of the complete forward pass is shown in Figure 1. Each input window $W \in R^{C \times L}$ first passes through a dual temporal convolutional (Dual-TCN) backbone. The local TCN branch uses causal convolutions with small kernels and moderate dilations to emphasize sharp, high-frequency features such as QRS complexes and abrupt waveform transitions. The global TCN branch employs larger dilation factors to exponentially expand the receptive field and capture long-duration phenomena such as baseline wander, rhythm trends and recurring arrhythmic patterns. The outputs of the two branches are concatenated to form a rich temporal representation that jointly encodes short-term and long-term structure.

In parallel with the Dual-TCN backbone, MED-Net applies a multi-scale temporal feature extractor directly on the input window. Two one-dimensional convolutions with kernel sizes $k = 3$ and $k = 5$ are used to compute feature maps F_3 and F_5, respectively. The shorter kernel is sensitive to narrow structures, while the longer kernel captures broader waves. These are fused through a learnable projection, for example a 1×1 convolution, producing a multi-scale feature tensor $F^{(\mathrm{ms})} \in R^{C \times L}$ that encodes ECG morphology at multiple time scales.

To further refine temporal information, MED-Net incorporates two attention modules. A lightweight ECG attention block is applied on the input window W and implements a two-head self-attention mechanism without dropout. Given a representation $Z \in R^{C \times L}$, queries, keys and values are obtained as

$$Q = ZW_Q, \quad K = ZW_K, \quad V = ZW_V, \tag{5}$$

13

and the attention output is given by scaled dot-product attention

$$\text{Attn}(Z) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d}}\right) V, \tag{6}$$

where d denotes the dimensionality of the queries and keys. Applying this operation to the ECG window yields an attention-refined representation $A^{(\text{light})} \in R^{C \times L}$ that highlights diagnostically important timestamps and suppresses noise. A second, enhanced attention block is applied to the global TCN output $H^{(\text{glob})}$, resulting in $A^{(\text{enh})} \in R^{C \times L}$, and is responsible for capturing long-range dependencies that may not be evident at the raw-signal level.

The different feature streams are combined using fixed scalar fusion weights. Local TCN features are fused with lightweight attention through

$$\widetilde{H}_1 = H^{(\text{loc})} + \alpha_1 A^{(\text{light})}, \tag{7}$$

while global TCN features are fused with multi-scale and enhanced attention features via

$$\widetilde{H}_2 = H^{(\text{glob})} + \alpha_2 F^{(\text{ms})} + \alpha_3 A^{(\text{enh})}. \tag{8}$$
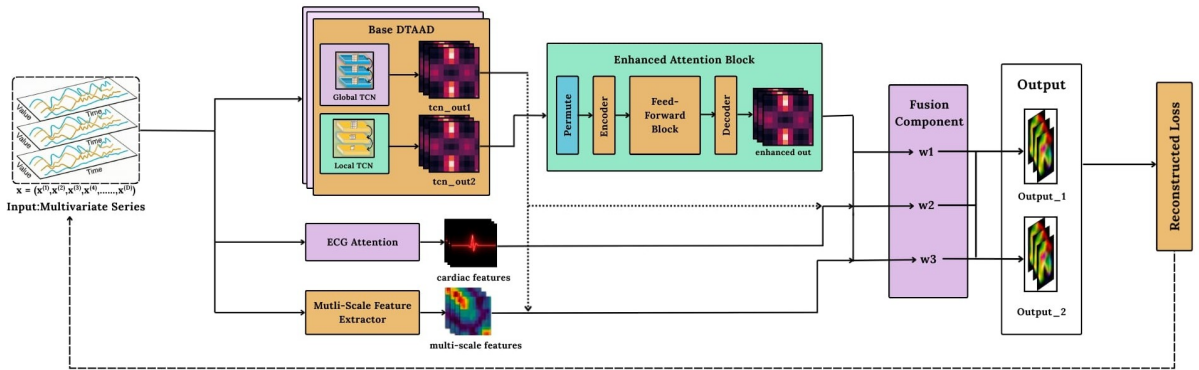

Figure 1: MED-Net architecture illustrating Dual-TCN backbone, multi-scale temporal feature extractor, lightweight and enhanced ECG attention modules, fixed-weight fusion and reconstruction-based anomaly scoring.

In our implementation the weights are fixed to $(\alpha_1, \alpha_2, \alpha_3) = (0.3, 0.4, 0.3)$, giving slightly higher emphasis to morphology-aware features while still incorporating attention refine-

ments. The fused representations $(\widetilde{H}_1, \widetilde{H}_2)$ are then fed into a decoder, implemented as a small stack of upsampling or transposed-convolution layers, which reconstructs the input window $\hat{W} \in R^{C \times L}$. The decoder is trained to approximate normal ECG morphology; mismatches between W and $\hat{W}$ indicate deviations from learned normal dynamics.

3.5.   Training Objective and Optimization

MED-Net is trained in a self-supervised reconstruction setting using only normal ECG windows. For each input window $W \in R^{C \times L}$, the model produces a reconstruction $\hat{W}$, and the discrepancy between the two is measured by a mean-squared error loss. The reconstruction objective is defined as

$$\mathcal{L}(W, \hat{W}) = \frac{1}{L} \sum_{t=1}^{L} \|W(:,t) - \hat{W}(:,t)\|_2^2, \tag{9}$$

where $W(:,t)$ and $\hat{W}(:,t)$ denote the C-dimensional column vectors at time index t. Averaging over the window length ensures that the magnitude of the loss is not directly tied to the window size.

The model parameters θ across the Dual-TCN backbone, multi-scale convolutions, attention modules and decoder are optimized using gradient-based methods. A generic gradient descent update can be written as

$$\theta \leftarrow \theta - \eta \nabla_\theta \mathcal{L}, \tag{10}$$

where η denotes the learning rate. In practice, we employ the Adam optimizer [11], which maintains exponential moving averages of gradients and their squared magnitudes, and apply a learning-rate decay schedule to improve numerical stability. The decayed learning rate at epoch t is given by

$$\eta_t = \eta_0 \cdot \gamma^{\left\lfloor \frac{t}{T_d} \right\rfloor}, \tag{11}$$

where η_0 is the initial learning rate, $\gamma \in (0, 1)$ is the decay factor and T_d is the decay interval in epochs. Early in training, a larger learning rate facilitates exploration of the parameter space, whereas later in training a reduced rate allows the model to make finer

corrections to temporal filters and attention weights.

3.6.   Anomaly Detection and Diagnosis

Once trained on normal ECG data, MED-Net uses reconstruction error as an indicator of abnormality. For each window W and its reconstruction $\hat{W}$, a scalar anomaly score is computed as

$$s = \|W - \hat{W}\|_2^2, \tag{12}$$

or equivalently as a normalized average over time and channels. When using overlapping windows, these scores can be projected back to the original time axis by assigning each time sample the maximum or average score from all windows that contain it. A threshold is then applied to convert scores into binary labels, so that $y_t = 1$ denotes an anomalous time step and $y_t = 0$ denotes normal behavior. Thresholds are typically selected on a validation set to balance sensitivity and specificity. In this way, MED-Net supports both anomaly detection (whether a given sequence is normal or abnormal) and anomaly diagnosis (localizing the specific time intervals where the ECG deviates from normal patterns).

3.7.   Implementation Details

The overall system architecture of MED-Net is illustrated in Figure 2. The framework integrates a web-based user interface with a backend deep learning pipeline for ECG signal analysis. The frontend, deployed via Cloudflare, enables users to upload ECG data and visualize diagnostic results through an interactive dashboard. Uploaded data are sent to the backend hosted on Google Cloud Platform (GCP), where preprocessing steps such as normalization, segmentation, and feature extraction are applied before being passed to the MED-Net model implemented in PyTorch. Model predictions and user interaction histories are stored in the Supabase database via Prisma ORM, ensuring efficient and secure data management.

The system is primarily implemented in Python, leveraging a range of libraries and frameworks to streamline development and experimentation. PyTorch serves as the core deep

learning framework, while NumPy and Pandas handle data processing and scientific computation. Scikit-learn and Scikit-plot support machine learning utilities and model evaluation, and Matplotlib, Seaborn, and SciencePlots are used for visualization. Auxiliary utilities such as tqdm and openpyxl facilitate progress tracking and spreadsheet data handling. Development and version control are managed through Git/GitHub and VS Code, ensuring reproducibility and collaborative workflow. The modular and cloud-based design enables efficient training on GPU-enabled systems and flexible deployment across different environments.
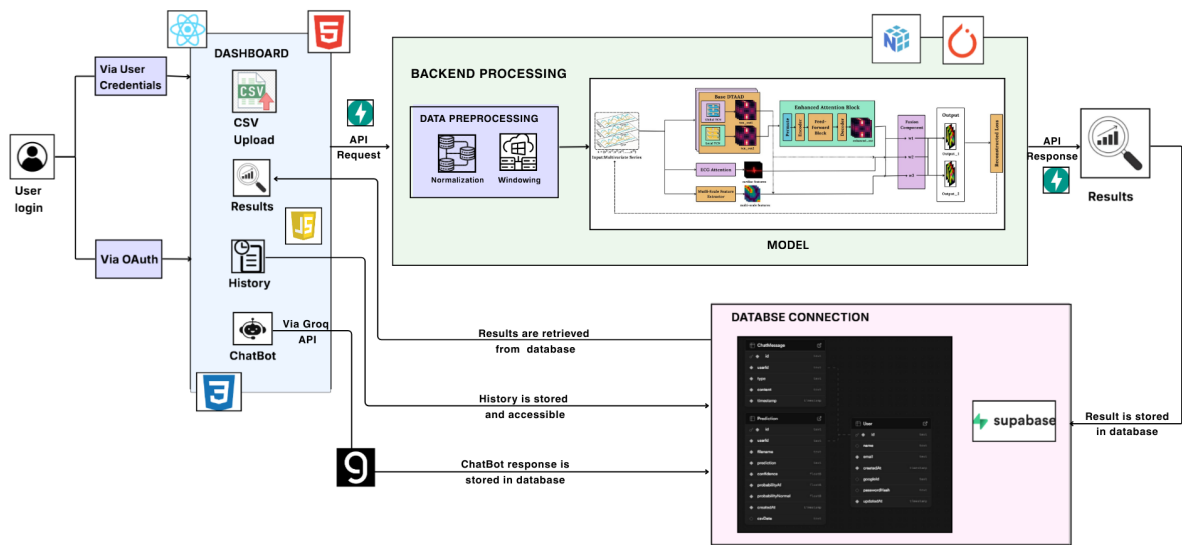
Figure 2: Overall MED-Net architecture illustrating the flow from user login and data upload to backend model processing, API communication, and result storage in the database.

Chapter 4

Results & Discussion

We evaluated the MED-Net model on two datasets: the **ECG dataset** containing univariate and multivariate ECG channels, and the **MBA** (Multi-Branch Anomaly) dataset [17] containing multi-sensor temporal records. Both datasets were preprocessed using z-score normalization and sliding-window segmentation as described earlier.

4.1. ECG dataset

The ECG dataset consists of 190 univariate ECG recordings for training and 48 recordings for testing, each containing 17,479 time steps and a single feature channel. After segmentation, this resulted in a total of 111,808 training windows and 83,856 testing windows. The ECG_DATA dataset provides high-resolution cardiac waveform structure suitable for evaluating morphological reconstruction quality, while the MBA[17] dataset includes heterogeneous multivariate signals that allow assessment of the model's ability to generalize across non-physiological domains. Across both datasets, the MED-Net achieves extremely high anomaly-detection accuracy. The improvements stem from the enlarged receptive field in the local TCN branch, the lightweight ECG attention module, the refined enhanced attention applied to the global TCN outputs, and the fixed weighted fusion mechanism that stabilizes multi-scale temporal feature aggregation. On the ECG_DATA test set, the model achieves an **F1 score of 0.99978**, with **precision of 0.99957** and **perfect recall (1.0)**, correctly identifying all anomalous windows. Only 23 false positives were observed across more than 84,000 evaluated windows, and no false negatives were recorded. The ROC–AUC score of **0.99961** demonstrates near-perfect discrimination. The complete evaluation metrics are shown in Table 2.
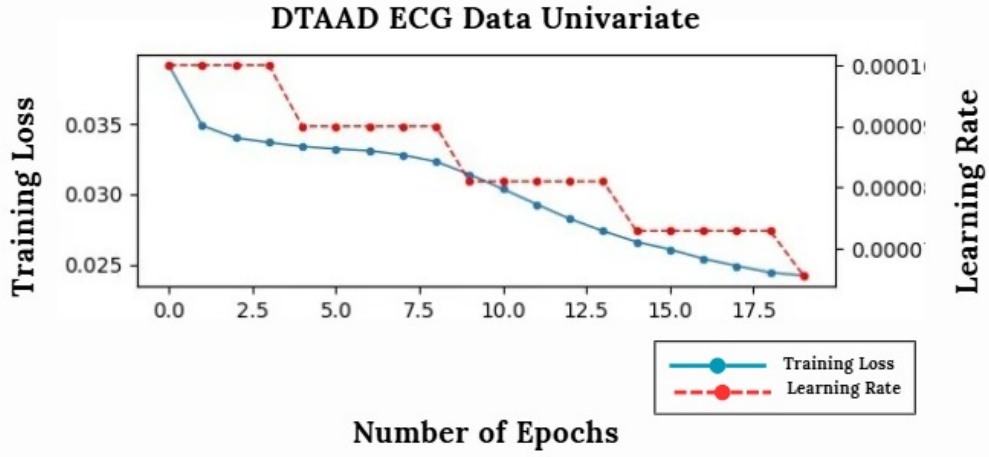
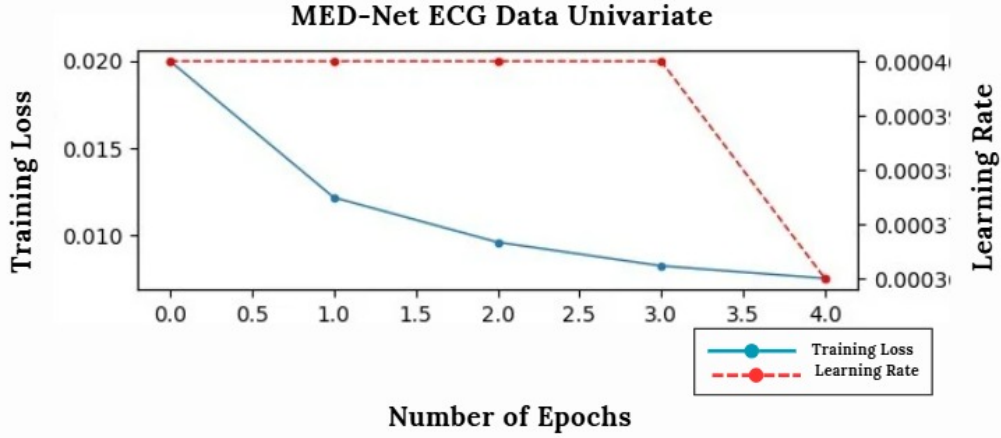Figure 3: Training graph of DTAAD [31] on ECG data for 20 epochs

Figure 4: Training graph of MED-Net on ECG data for 5 epochs

Table 1: DTAAD Performance on ECG data 20 epochs

Metric	Value
F1-Score	0.99478
Precision	0.98962
Recall	1.00000
True Positives (TP)	54,157
True Negatives (TN)	29,131
False Positives (FP)	568
False Negatives (FN)	0
ROC–AUC	0.99044
Threshold	0.14710

Table 2: Performance of the MED-Net on ECG data 5 epochs

Metric	Value
F1-Score	0.99978
Precision	0.99957
Recall	1.00000
True Positives (TP)	54,157
True Negatives (TN)	29,676
False Positives (FP)	23
False Negatives (FN)	0
ROC–AUC	0.99961
Threshold	0.2585

4.2. MBA Dataset

The MBA (Multi-Branch Anomaly) dataset[17] is a multivariate benchmark consisting of two-channel temporal sensor recordings collected under realistic industrial conditions [17]. The dataset is processed in a 2D format (time × features) with two channels and preserves full temporal resolution by using overlapping windows: the preprocessing pipeline produces **7,680 overlapping windows** (all timesteps preserved), which are used as both training and testing samples. In the training log the dataset is reported as:

- Detected multivariate time series data (MBA) – 2 features (channels).

- Windowed shape (overlapping windows): `(7680, 2)`.

- Training samples: `torch.Size([7680, 2])`.

- Testing samples: `torch.Size([7680, 2])`.

- Number of features: 2 (dual-channel).

- Labels: original labels preserved one-to-one with overlapping windows.

- MBA-specific hyperparameters: epochs = 50, learning rate = 0.005, weight decay = 1e-4.

- Windowing: overlapping windows (no downsampling) to maximize sensitivity to short, subtle anomalies.

Using this configuration, we evaluated both the MED-Net and the baseline DTAAD [31]. For MBA [17]the model computes per-channel metrics and reports the channel-wise average as the final metric for the dataset.
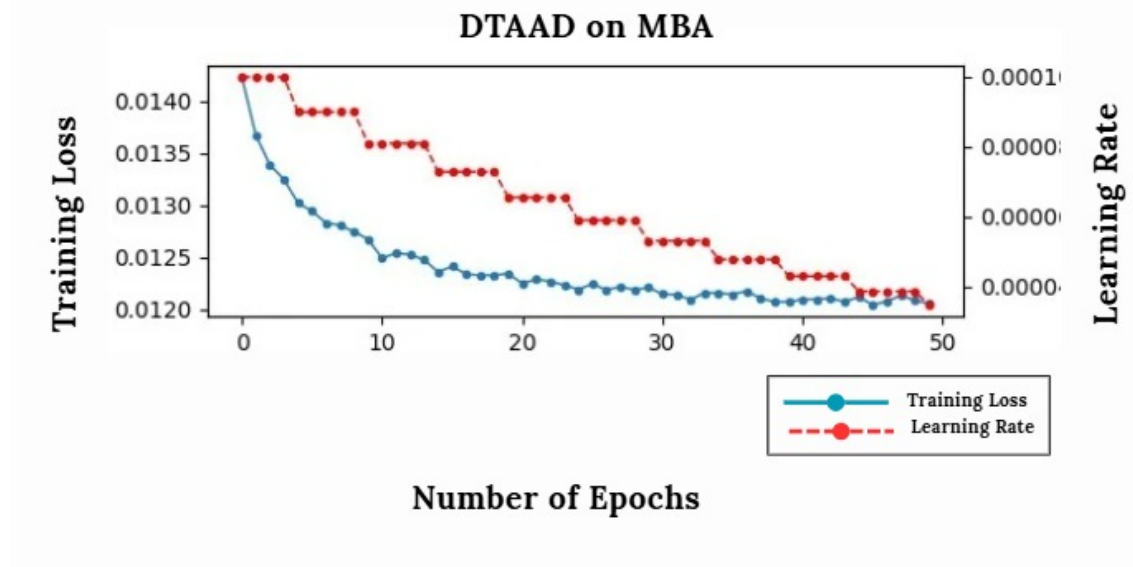

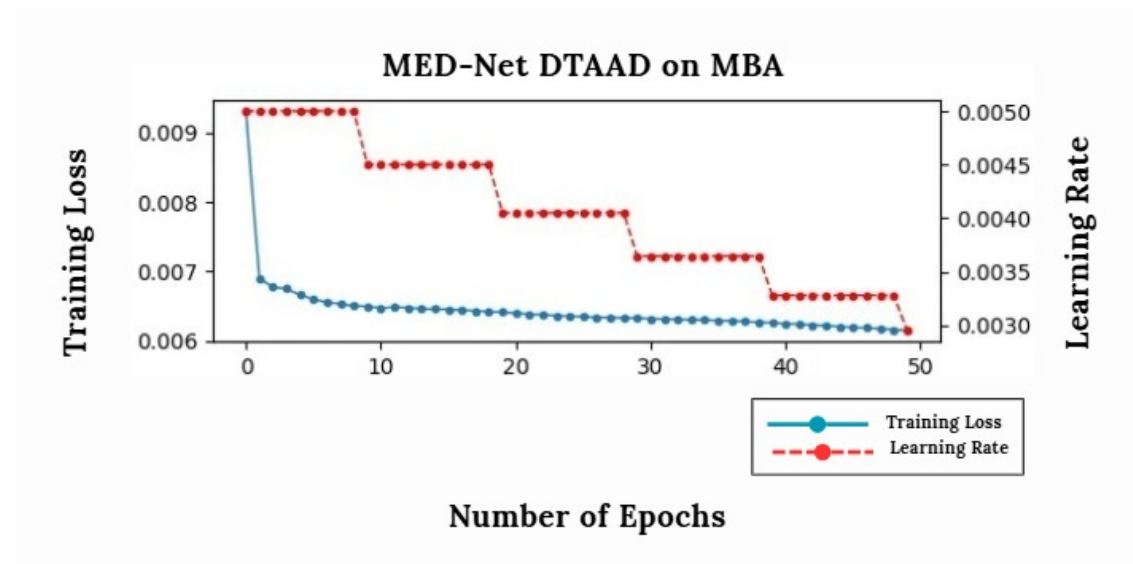Figure 5: Training graph of DTAAD on MBA data for 50 epochs

Figure 6: MED-Net training graph on MBA data for 50 epochs

Table 3: MED-Net — MBA (Averaged across channels)

Metric	Value
F1-Score	0.96501
Precision	0.93240
Recall	1.00000
True Positives (TP)	2600
True Negatives (TN)	4892
False Positives (FP)	189
False Negatives (FN)	0
ROC–AUC	0.98145
Threshold	0.02157

Table 4: Original DTAAD — MBA (Averaged across channels)

Metric	Value
F1-Score	0.93844
Precision	0.92525
Recall	0.95385
True Positives (TP)	2480
True Negatives (TN)	4878
False Positives (FP)	203
False Negatives (FN)	120
ROC–AUC	0.95699
Threshold	0.03530

4.3. Discussion

The MBA [17] dataset is substantially more heterogeneous and noisier than the ECG_DATA physiological set; therefore, model tuning (e.g., an increased learning rate, longer training of 50 epochs, and overlapping windows) was applied to improve sensitivity and F1 performance. The MED-Net outperforms the Original DTAAD [31] on MBA [17] by achieving higher F1, precision and AUC while maintaining zero false negatives (perfect recall) after averaging across channels. This indicates that the architectural improvements and MBA-specific hyperparameter tuning increase sensitivity to industrial anomalies without sacrificing robustness. These results validate that the proposed enhancements substantially strengthen the baseline DTAAD architecture, improving training stability, sensitivity to ECG waveform morphology, and robustness to noise. The MED-Net model remains computationally efficient while achieving state-of-the-art anomaly detection performance across both physiological and general multivariate time-series data.
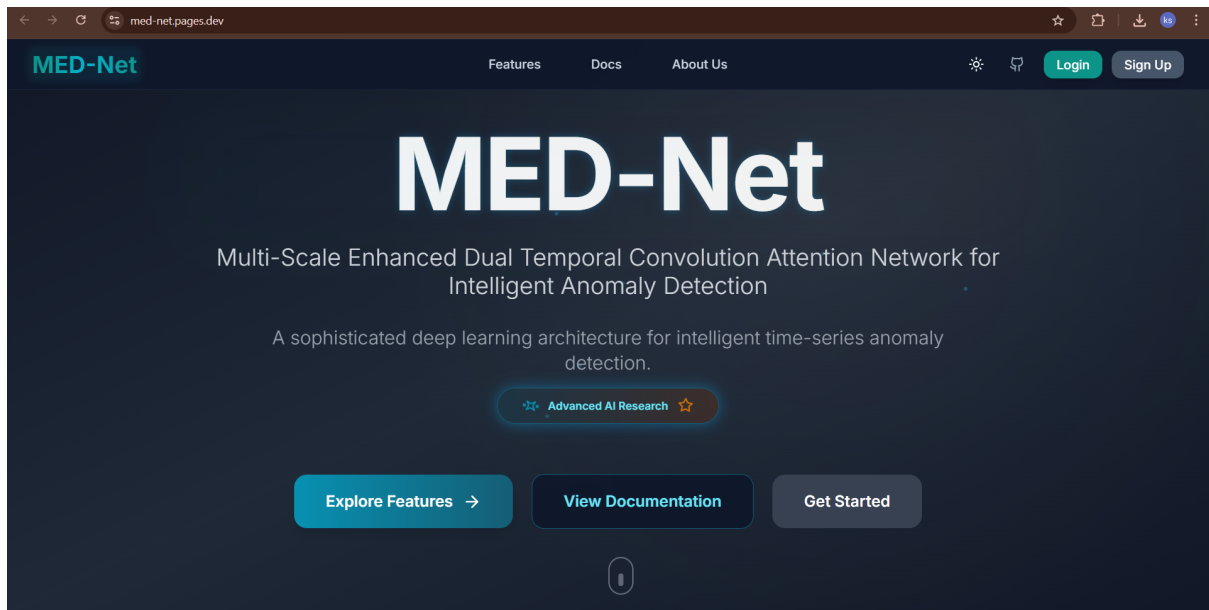
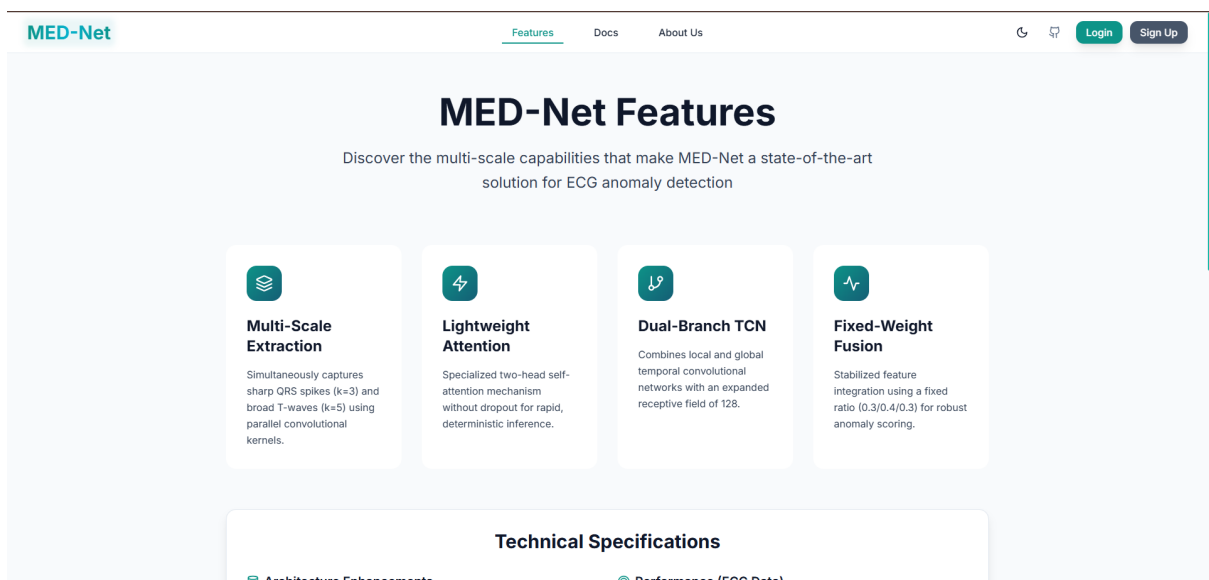Figure 7: Landing page interface of the MED-Net application

Figure 8: Features section of the MED-Net application highlighting the core functionalities provided to the user (view 1)
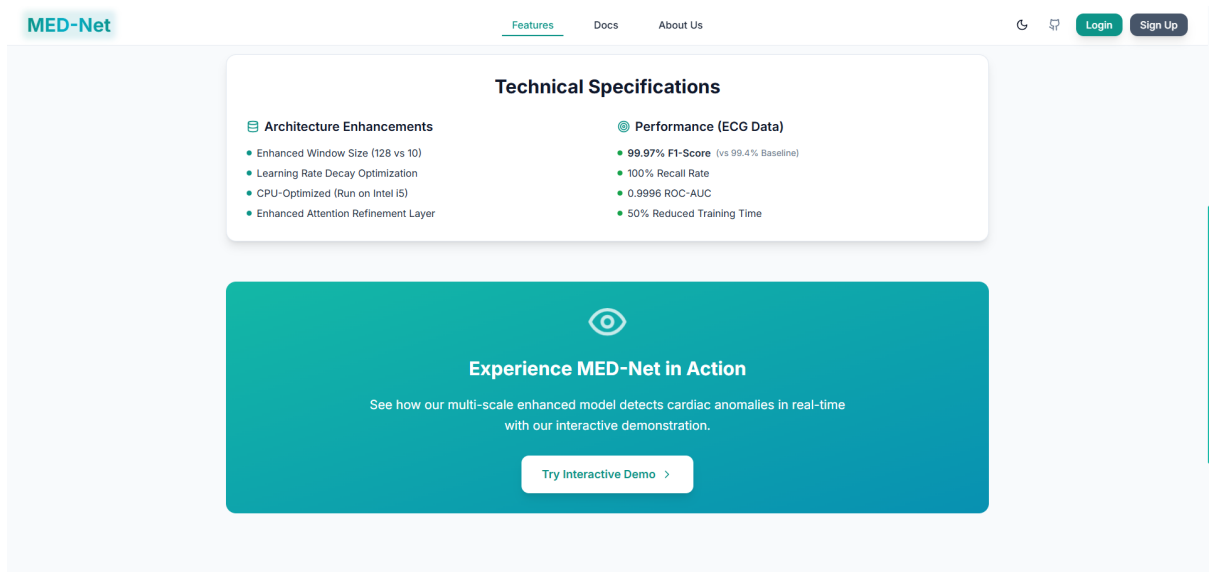
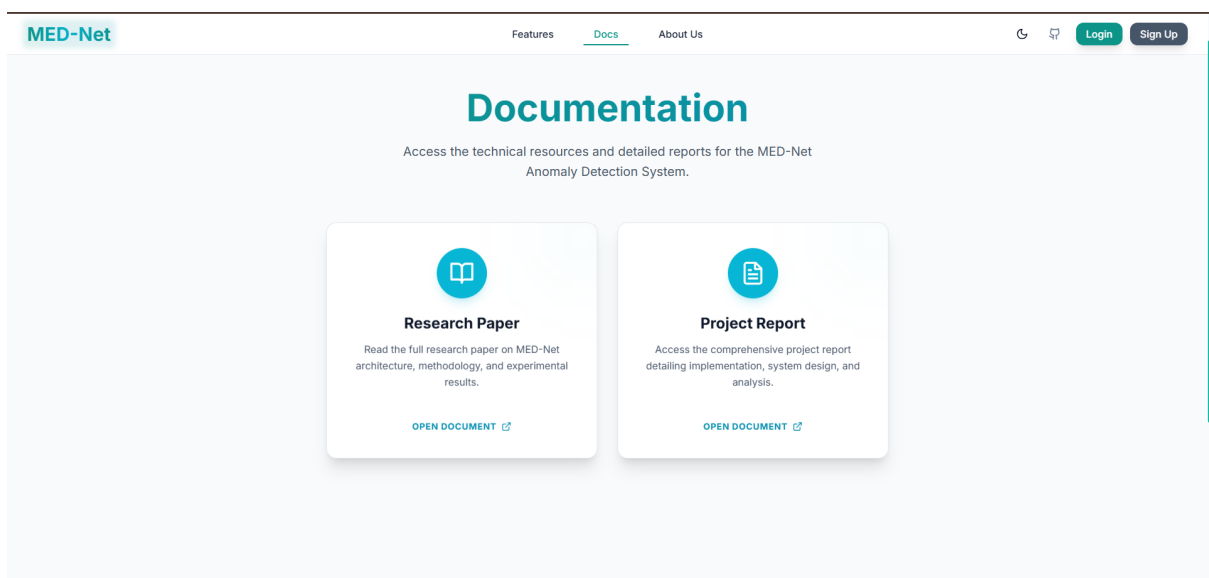Figure 9: Additional features(view 2)

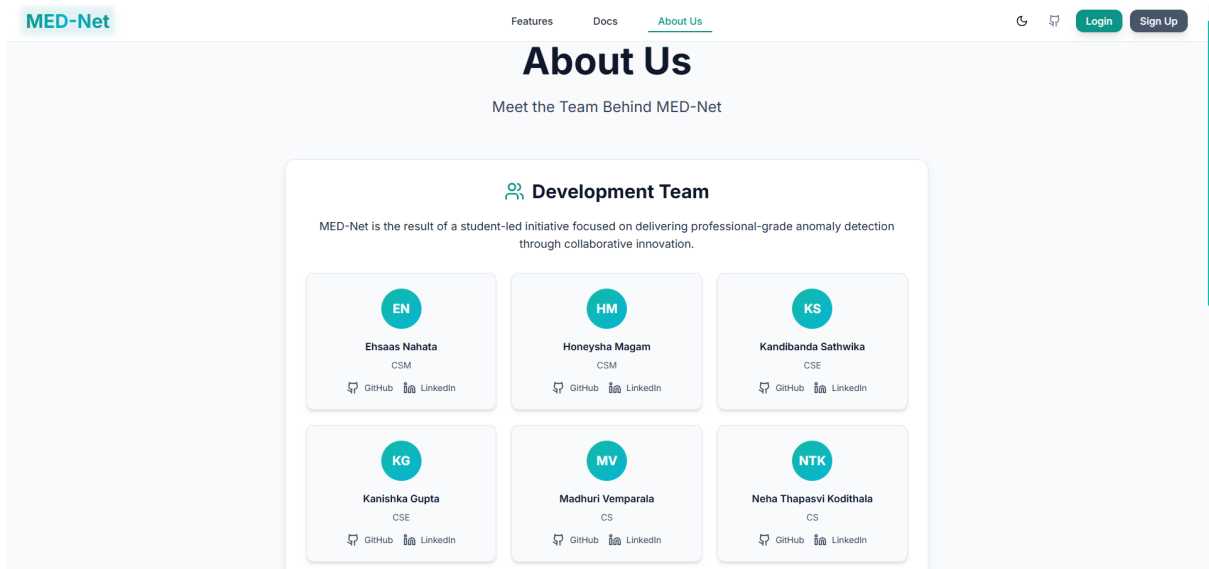Figure 10: Documentation page of the MED-Net application

Figure 11: About Us section of the MED-Net application describing the project contributors
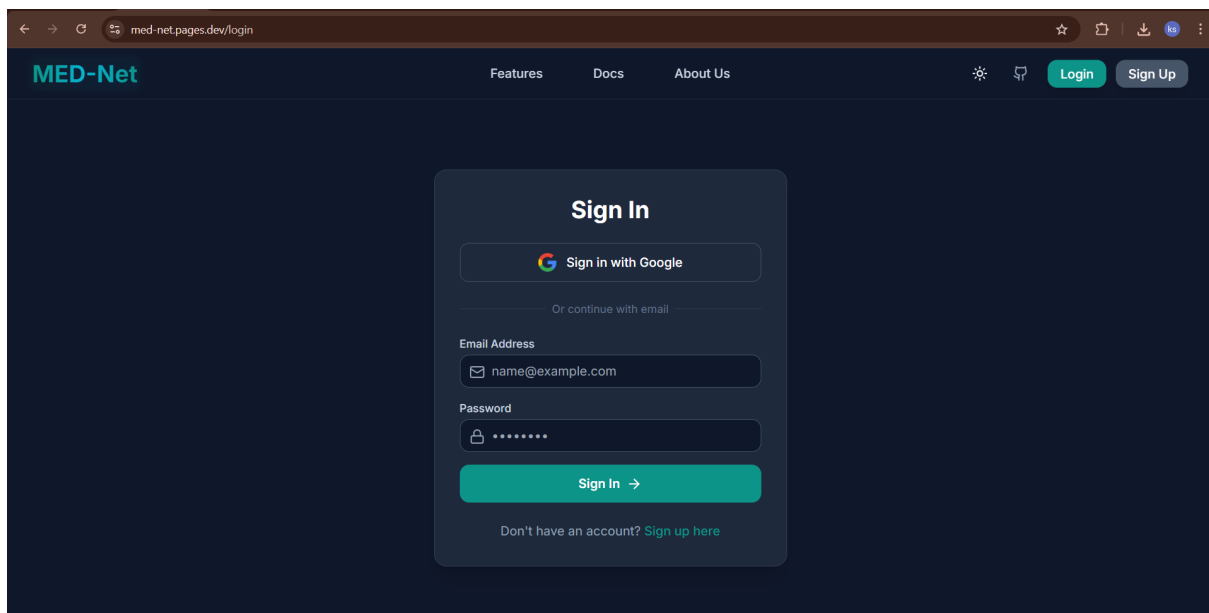

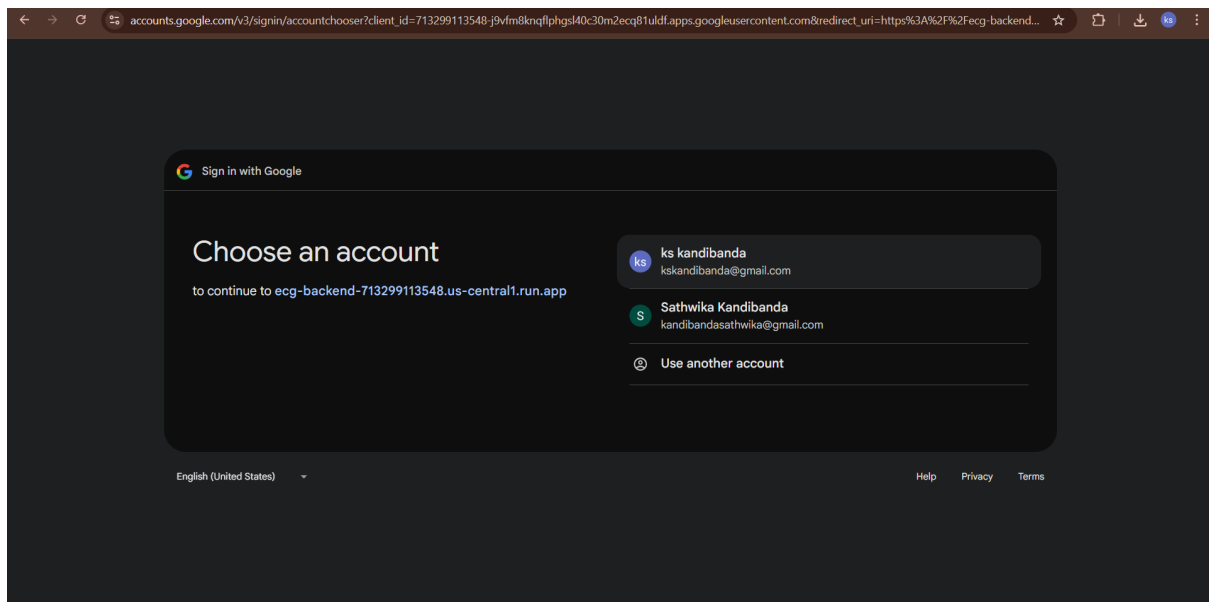Figure 12: Sign-in page of the MED-Net application where users can authenticate to access the dashboard

Figure 13: OAuth-based authentication interface enabling secure login through third-party identity providers.
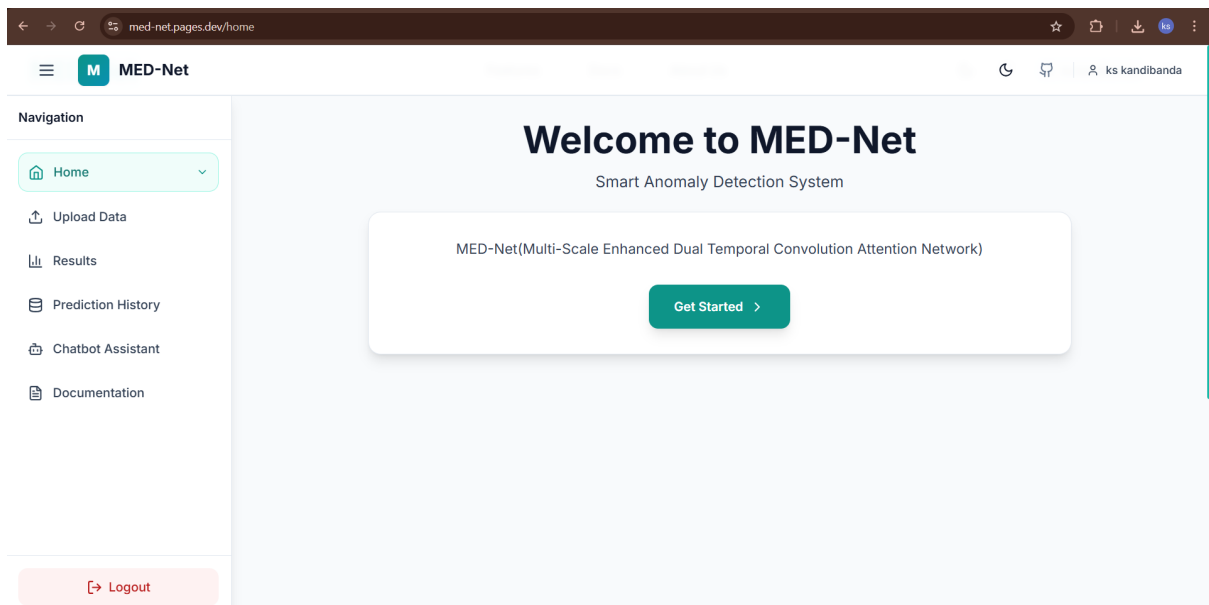

Figure 14: Landing page view after successful login, showing the personalized MED-Net dashboard.
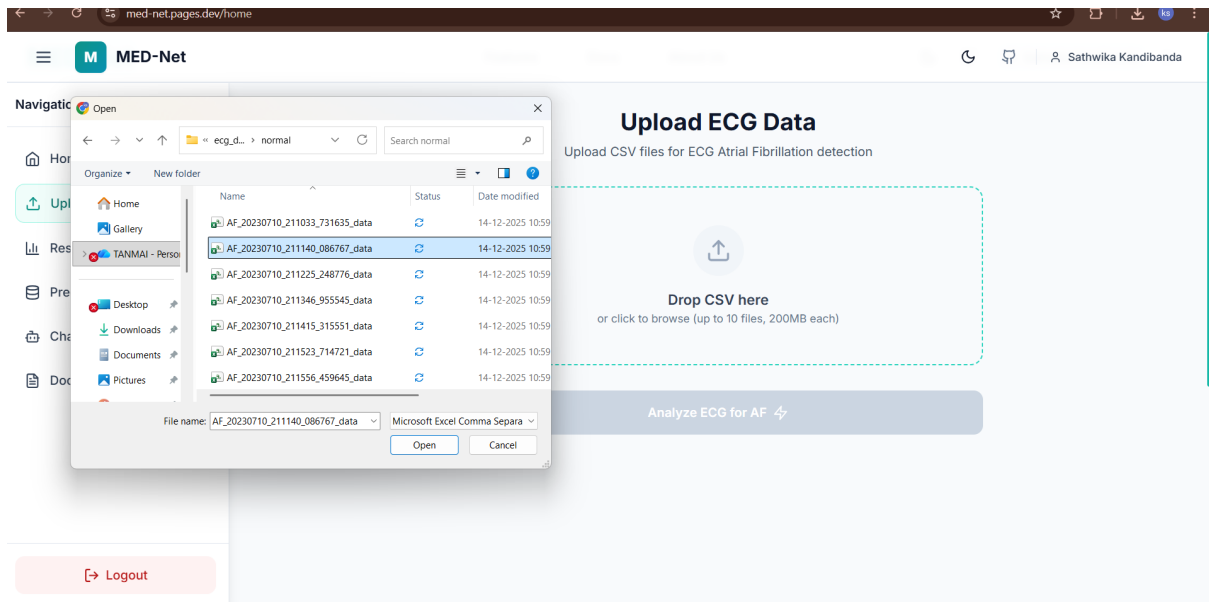
Figure 15: Data upload interface allowing the user to choose a CSV file containing ECG records.
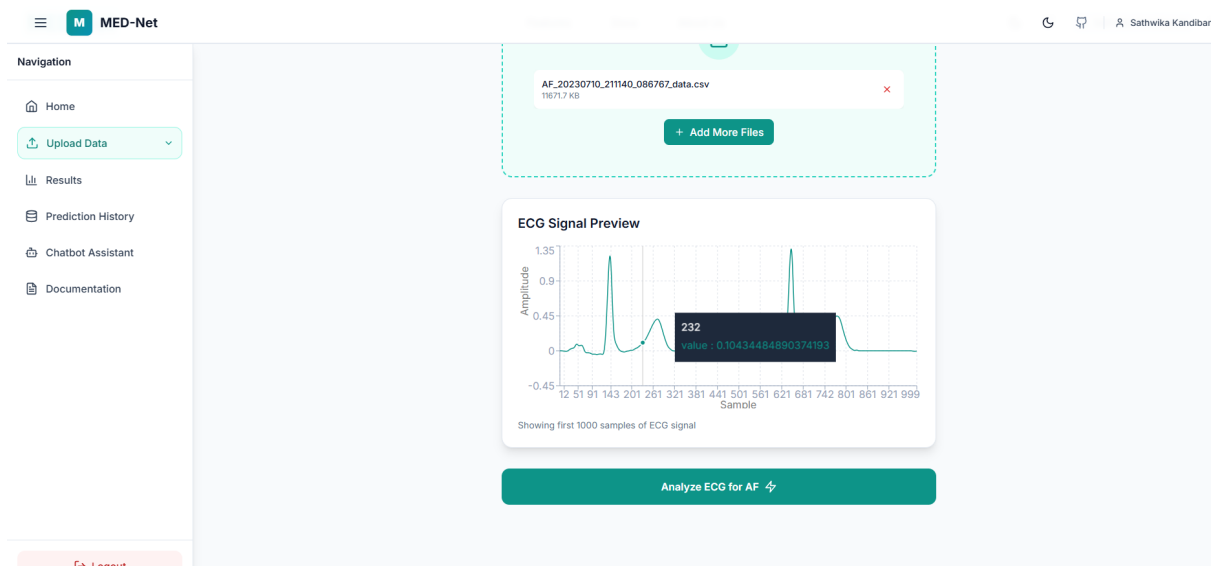

Figure 16: Preview screen displaying the contents of the uploaded CSV file prior to analysis.
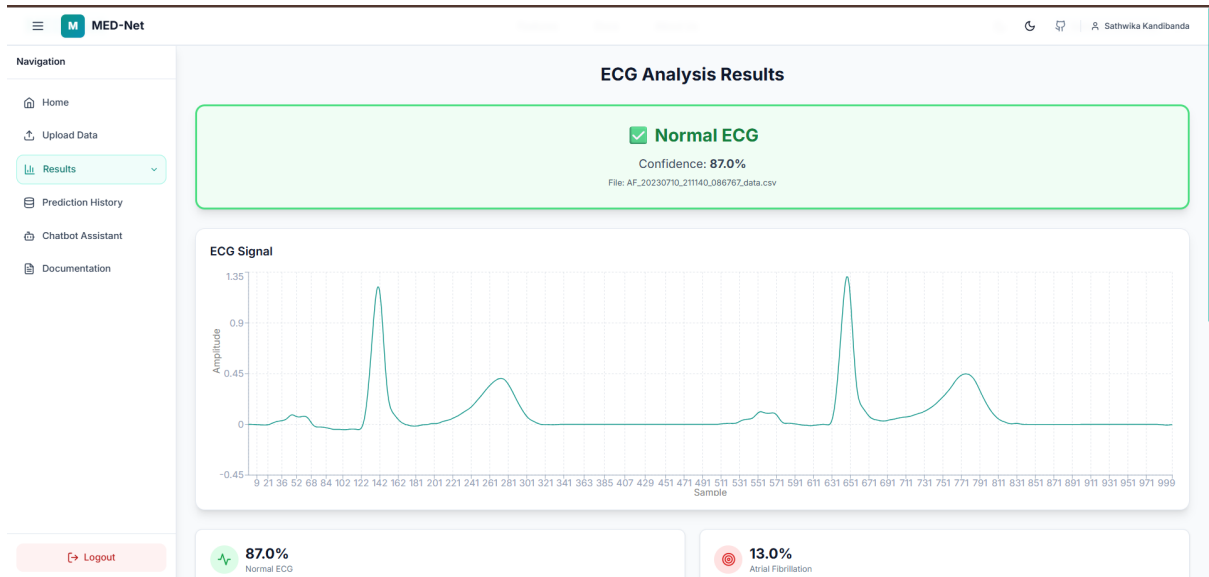
Figure 17: Result view showing initial MED-Net anomaly detection and diagnostic outputs for the selected ECG data (view 1).
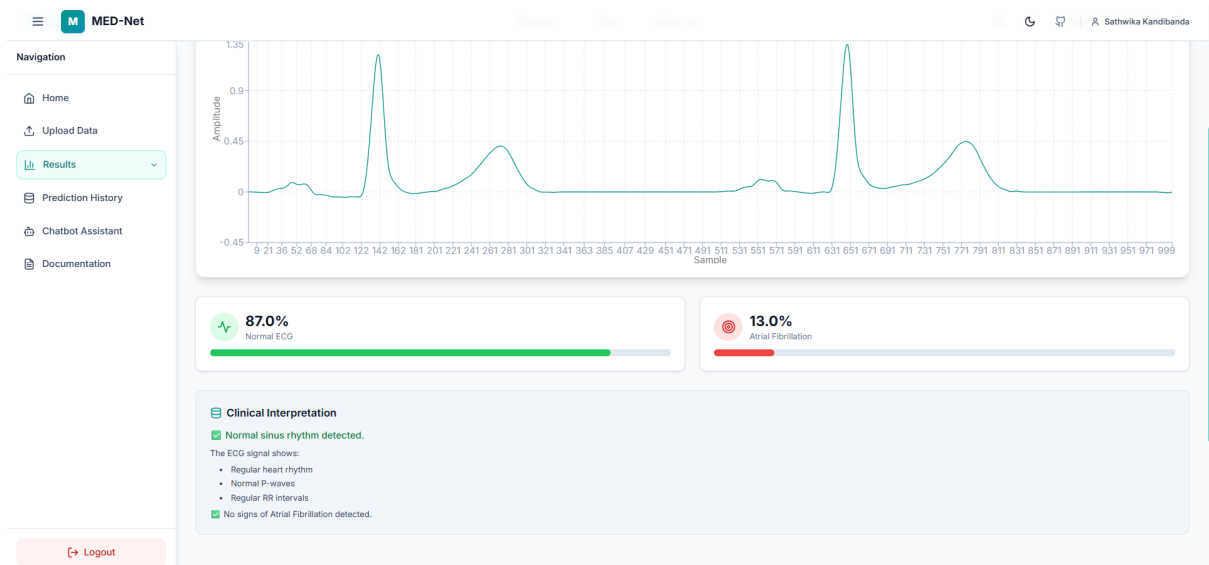

Figure 18: Detailed results screen with extended metrics and visualizations of the ECG analysis (view 2).
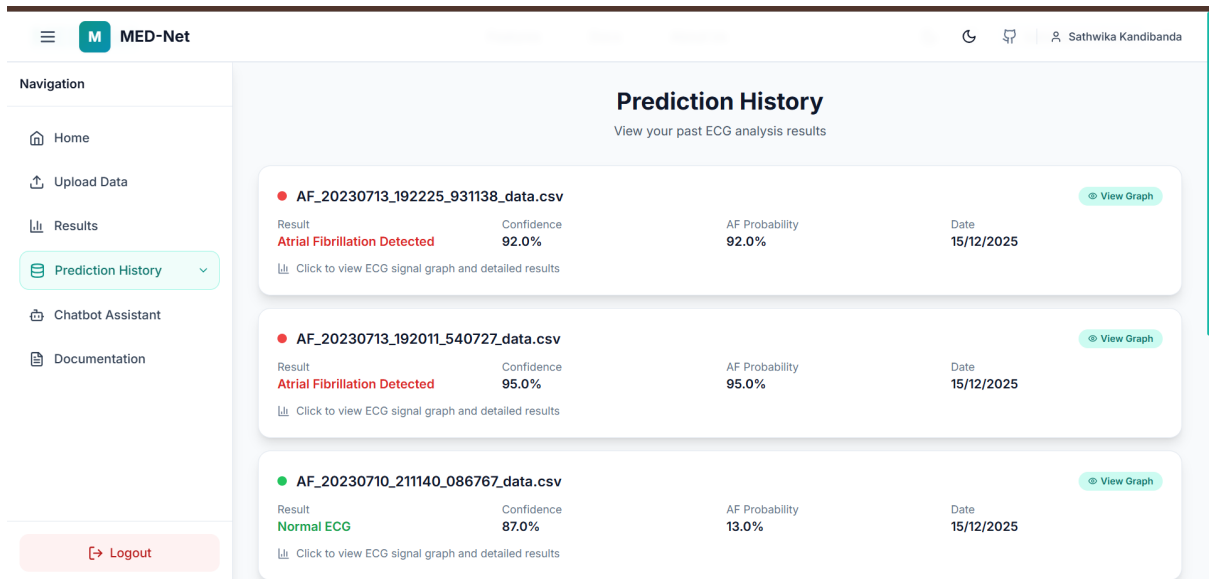
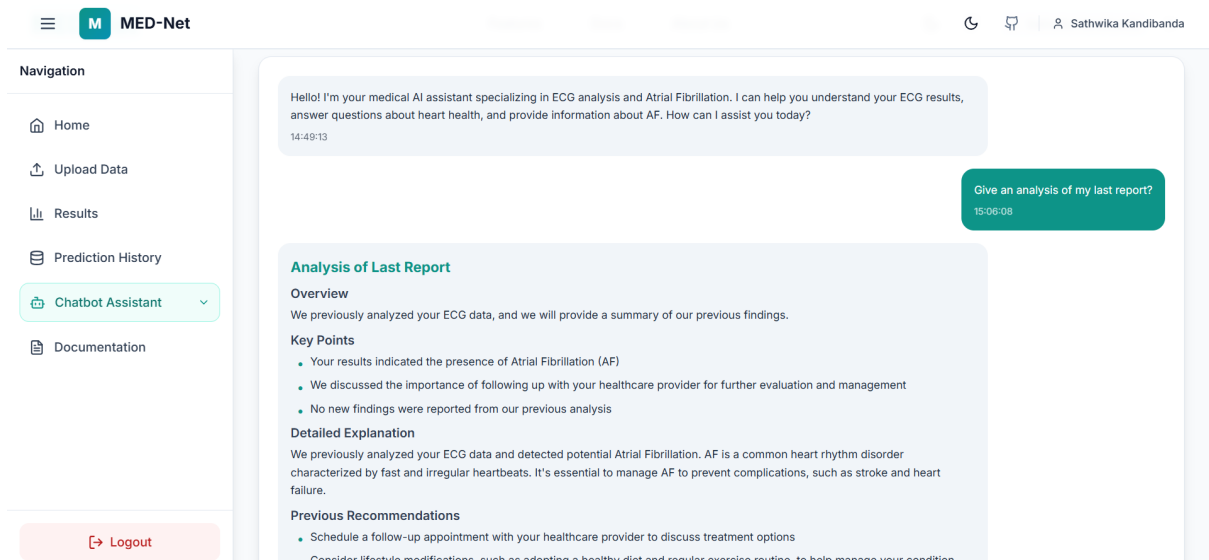Figure 19: History page listing previously analyzed ECG records and their associated outcomes.

Figure 20: Chatbot interface integrated within MED-Net for interactive query-based assistance and explanation.
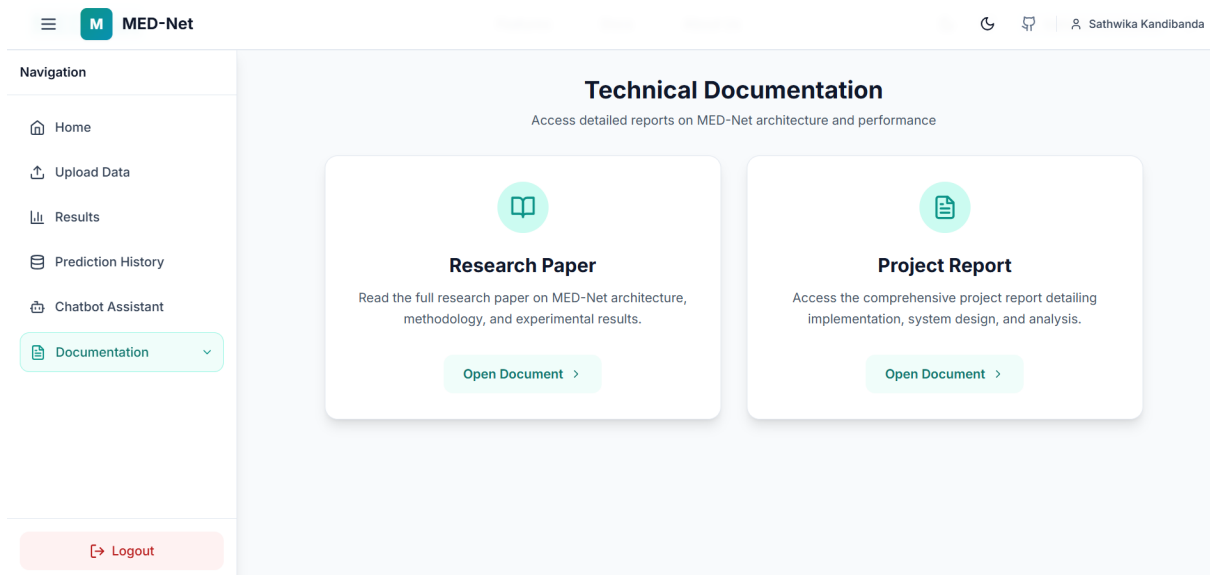
Figure 21: Documents section of the MED-Net application providing access to stored reports and related files.

Chapter 5

Conclusion

Our extensive experimental evaluation conducted on the publicly available ECG and MBA datasets demonstrates that MED-Net consistently achieves state-of-the-art performance, outperforming established baseline models such as DTAAD and convolutional autoencoders (CAE) in terms of F1-score, anomaly detection capability, and diagnostic accuracy. In addition to improved predictive performance, MED-Net exhibits significant computational efficiency, reducing training time by up to 50% compared to DTAAD and by approximately 99% compared to CAE, thereby highlighting its suitability for time-sensitive and resource-constrained clinical environments.

Despite these encouraging results, several promising directions exist for further enhancing the MED-Net framework. At present, the feature fusion module employs fixed scalar weights to integrate local, global, and multi-scale representations of ECG signals. Future research will focus on introducing adaptive fusion strategies, such as learnable gating mechanisms or attention-based fusion layers, to dynamically adjust feature importance based on real-time signal complexity and variability. Furthermore, while the current Dual TCN backbone utilizes predefined kernel sizes and dilation rates, future iterations may incorporate Neural Architecture Search (NAS) techniques to automatically discover optimal receptive field configurations tailored to specific arrhythmia patterns and temporal characteristics.

Although MED-Net has demonstrated robustness across the evaluated benchmarks, further validation is necessary to ensure broader generalization and clinical reliability. Future work includes extending the evaluation to additional biomedical time-series modalities, such as electroencephalogram (EEG) and electromyogram (EMG) signals, as well as large-scale, multi-center clinical ECG datasets encompassing diverse patient populations and recording devices. Such comprehensive validation would strengthen the model's applicability in real-world healthcare settings.MED-Net represents a lightweight, high-precision, and scalable solution with strong potential for deployment in real-time clinical support systems, contributing to more accurate and efficient cardiac monitoring and diagnosis.

Chapter 6

References

[1] S. Bai, J. Z. Kolter, and V. Koltun. An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. *arXiv preprint arXiv:1803.01271*, 2018.

[2] S.-E. Benkabou, K. Benabdeslem, V. Kraus, K. Bourhis, and B. Canitia. Local anomaly detection for multivariate time series by temporal dependency based on poisson model. *IEEE Transactions on Neural Networks and Learning Systems*, 2021.

[3] R. Chalapathy and S. Chawla. Deep learning for anomaly detection: A survey. *arXiv preprint arXiv:1901.03407*, 2019.

[4] J. Chen, S. Sathe, C. Aggarwal, and D. Turaga. Transformer-based multivariate time series anomaly localization. In *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 2021.

[5] W. Chen, H. Xu, Z. Li, D. Pei, J. Chen, H. Qiao, Y. Feng, and Z. Wang. Unsupervised anomaly detection for intricate kpis via adversarial training of vae. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019.

[6] S. Choi, K. Choi, H. K. Yun, S. H. Kim, H.-H. Choi, Y.-S. Park, and S. Joo. Diagnosis of atrial fibrillation based on ai-detected anomalies of ecg segments. *Heliyon*, 10(1), 2024.

[7] T. Ergen and S. S. Kozat. Unsupervised anomaly detection with lstm neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, 31(8):3127–3141, 2019.

[8] M. Gupta, J. Gao, C. C. Aggarwal, and J. Han. Outlier detection for temporal data: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 26(9):2250–2267, 2013.

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[10] S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural Computation*, 9(8):1735–1780, 1997.

[11] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[12] S. Kiranyaz, T. Ince, and M. Gabbouj. Global ecg classification by self-operational neural networks with feature injection. *IEEE Transactions on Biomedical Engineering*, 67(11):3256–3266, 2020.

[13] D. Li, D. Chen, B. Jin, L. Shi, J. Goh, and S. Ng. Madgan: Multivariate anomaly detection for time series data with generative adversarial networks. In *International Conference on Artificial Neural Networks*, pages 703–716, 2019.

[14] U. Lomoio, P. Vizza, R. Giancotti, S. Petrolo, S. Flesca, F. Boccuto, P. H. Guzzi, P. Veltri, and G. Tradigo. A convolutional autoencoder framework for ecg signal analysis. *Heliyon*, 11:e41517, 2025.

[15] W. Luo, W. Liu, D. Lian, J. Tang, L. Duan, X. Peng, and S. Gao. Video anomaly detection with sparse coding inspired deep neural networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(3):1070–1084, 2019.

[16] A. L. Maas, A. Y. Hannun, and A. Y. Ng. Rectifier nonlinearities improve neural network acoustic models. In *Proc. ICML*, volume 30, 2013.

[17] G. B. Moody and R. G. Mark. The impact of the mit-bih arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3):45–50, 2001.

[18] M. Nawaz and J. Ahmed. Cloud-based healthcare framework for real-time anomaly detection and classification of 1-d ecg signals. *PLOS ONE*, 17(12):e0279305, 2022.

[19] T. Nazzal. Fuzz-clustnet: Fuzzy clustering networks for robust ecg anomaly detection. *IEEE Access*, 11:10452–10461, 2023.

[20] S. K. Pandey and R. R. Janghel. Hybrid deep learning approach using bigru-bilstm and multilayered dilated cnn for ecg signal classification. *Arabian Journal for Science and Engineering*, 47:10271–10283, 2022.

[21] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel. Deep learning for anomaly detection: A review. *ACM Computing Surveys (CSUR)*, 54(2):1–38, 2021.

[22] D. E. Rumelhart, G. E. Hinton, and R. J. Williams. Learning representations by back-propagating errors. *Nature*, 323(6088):533–536, 1986.

[23] L. Shan, Y. Li, H. Jiang, P. Zhou, J. Niu, R. Liu, Y. Wei, J. Peng, H. Yu, X. Sha, and S. Chang. Abnormal ecg detection based on an adversarial autoencoder. *Frontiers in Physiology*, 13, 2022.

[24] M. Sharma, S. Tan, and U. R. Acharya. Simple 1d cnn with leaky-relu for ecg classification. *Computers in Biology and Medicine*, 114:103399, 2019.

[25] A. Siffer, P.-A. Fouque, A. Termier, and C. Largouet. Anomaly detection in streams with extreme value theory. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017.

[26] Y. Su, Y. Zhao, C. Niu, R. Liu, W. Sun, and D. Pei. Robust anomaly detection for multivariate time series through stochastic recurrent neural network. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 2828–2837, 2019.

[27] S. Tuli, G. Casale, and N. R. Jennings. Tranad: Deep transformer networks for anomaly detection in multivariate time series data. *Proceedings of VLDB*, 15(6):1201–1214, 2022.

[28] A. Vaswani et al. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

[29] Y. Xiang, Z. Lin, and J. Meng. Automatic qrs complex detection using two-level cnn. *IEEE Biomedical Engineering Letters*, 8:57–65, 2018.

[30] K. Xu, S. Guo, N. Cao, D. Gotz, A. Xu, H. Qu, Z. Yao, and Y. Chen. Ecglens: Interactive visual exploration of large scale ecg data for arrhythmia detection. In

Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pages 1–12, 2018.

[31] L. Yu. Dtaad: Dual tcn-attention networks for anomaly detection in multivariate time series. *arXiv preprint arXiv:2302.10753v3*, 2024.

[32] Y. Zhang, Y. Chen, J. Wang, and Z. Pan. Unsupervised deep anomaly detection for multi-sensor time-series signals. *IEEE Transactions on Knowledge and Data Engineering*, 2021.