World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(Research Article)

Light weight neural network for ECG and EEG anomaly detection in IOT edge sensors

Sneha Pujari *

E and TC Department, GCOE Yavatmal, Maharashtra, India.

World Journal of Advanced Engineering Technology and Sciences, 2024, 11(02), 269–280

Publication history: Received on 15 February 2024; revised on 27 March 2024; accepted on 29 March 2024

Article DOI: https://doi.org/10.30574/wjaets.2024.11.2.0111

Abstract

Lightweight neural network designed for detecting anomalies in Electrocardiogram (ECG) and Electroencephalogram (EEG) signals at IoT edge sensors. By optimizing neural network architectures, we achieve high accuracy in anomaly detection while minimizing computational demands and memory usage. Experimental results validate the effectiveness of our approach in real-world scenarios, promising improved healthcare monitoring with early detection of abnormal ECG and EEG patterns at the edge.

Keywords: Electrocardiagram (ECG); Electroencephalogram (EEG); Anomaly detection; Lightweight neural network.

1. Introduction

Anomaly detection in ECG and EEG signals, emphasizing the importance of timely interventions in healthcare settings. The motivation could stem from challenges faced by traditional methods, such as computational constraints, and the potential of edge computing to address these challenges.

Objectives may include the development of a novel lightweight neural network architecture, achieving high accuracy in anomaly detection while minimizing computational demands, and evaluating the effectiveness of the proposed approach in scenarios.

Contributions may include the introduction of a novel lightweight neural network architecture tailored for ECG and EEG anomaly detection, the demonstration of its effectiveness through experimental evaluations, and its potential to enhance healthcare monitoring systems by enabling early detection of abnormal patterns at the edge.

Continuous monitoring of physiological signals such as ECG using IoT enabled wearable devices is widely considered a solution to mitigate the costs and healthcare risks associated with CVDs.

Several works attempt to classify ECG. Attempting multi-class classification on edge sensors may not be ideal and mostly redundant, due to the computational complexity involved. Typical sampling rate of ECG in real-world sensors is around 250Hz, while all the existing works use MIT-BIH records in its original, but rather unusual sampling rate of 360Hz for performance evaluation and this results in unrealistic, but higher performance. MIT-BIH records are highly imbalanced in terms of class distribution, and these are not carefully handled by much existing work a device, multi-class classification brings limited added value compared to simple anomaly detection.

* Corresponding author: Sneha Pujari.

1.1. Background

The Internet of Things (IoT) has witnessed exponential growth in recent years, revolutionizing various sectors including healthcare. IoT devices are increasingly being deployed to monitor patients' physiological signals, such as Electrocardiogram (ECG) and Electroencephalogram (EEG), in real-time. Edge computing, which involves processing data closer to its source rather than in centralized data centers, has emerged as a critical enabler for real-time analytics in IoT applications.

1.2. Motivation

Timely anomaly detection in ECG and EEG signals is crucial for identifying potential health issues and enabling prompt intervention. However, traditional anomaly detection methods often face challenges when deployed on resource-constrained IoT edge sensors. These methods are computationally intensive and may require significant memory and processing power, making them impractical for edge deployment.

1.3. Objectives

To address the computational constraints associated with anomaly detection in ECG and EEG signals at IoT edge sensors. We propose a novel lightweight neural network architecture optimized for efficient processing on resource-constrained devices. Our objectives include: Developing a lightweight neural network architecture specifically tailored for ECG and EEG anomaly detection. Minimizing computational demands and memory usage while maintaining high accuracy. Evaluating the effectiveness of the proposed approach through extensive experimentation on real-world datasets.

1.4. Contributions

The primary contributions of this research can be summarized as follows: Introduction of a novel lightweight neural network architecture for ECG and EEG anomaly detection in IoT edge sensors.

Demonstration of the effectiveness of the proposed approach in achieving high accuracy while minimizing computational overhead.

Validation of the approach through comprehensive experimentation on diverse datasets, highlighting its potential for real-world deployment in healthcare monitoring systems .

2. Related Work

2.1. Traditional Anomaly Detection Methods

Traditional anomaly detection methods often rely on statistical techniques, signal processing algorithms, or rule-based approaches.

While these methods can be effective in certain scenarios, they may struggle to handle the complexity and variability of physiological signals like ECG and EEG, especially when deployed on resource-constrained edge devices.

2.2. Edge Computing in Healthcare

Edge computing has gained significant traction in healthcare applications due to its ability to process data closer to the point of generation.

By leveraging edge computing, healthcare providers can achieve real-time analytics, reduce latency, and ensure data privacy and security.

2.3. Neural Network Approaches for ECG and EEG Analysis

Neural networks have shown promise in various healthcare applications, including ECG and EEG analysis.

Deep learning models, in particular, have demonstrated remarkable performance in tasks such as arrhythmia detection in ECG signals and seizure prediction in EEG recordings.

2.4. Lightweight Neural Networks

Lightweight neural networks have emerged as a research area aimed at developing efficient models that can run on resource-constrained devices.

These models typically involve techniques such as model compression, quantization, and architecture optimization to reduce memory and computational requirements while maintaining performance.

ECG is usually affected by various noises like baseline wander (low-frequency noise in the range of 0-0.3Hz).

electrode contact noise, motion artifacts, power line interference (PLI) etc. which affects the efficacy of signal analysis. Many real-world ECG devices perform baseline.

wander and PLI removal during acquisition and present a clean ECG signal at a typical sample rate of 250Hz.

To emulate this we perform the following data processing steps:-

- Discrete Wavelet Transform (DWT) based denoising
- PLI removal using a standard IIR notch filter at 60Hz
- Re-sampling from 360Hz to 250Hz.
- Illustration of ECG noise removal using the above steps is shown in Figure 1.

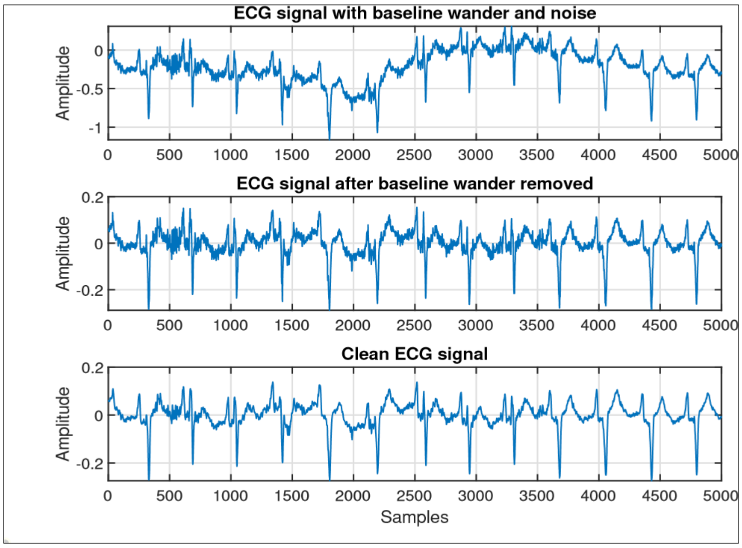

Figure 1 ECG Signal before and after denoising and notch filtering

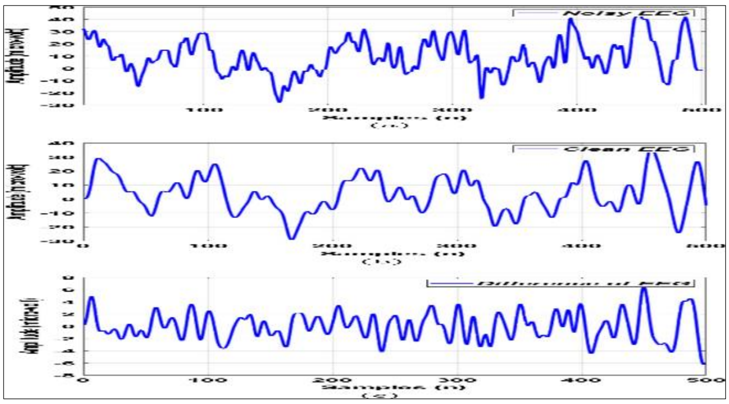

Figure 2 EEG Signal before and after denoising and notch filtering

3. Lightweight Neural Network Architecture

3.1. Data Preprocessing

The raw ECG and EEG signals are preprocessed to remove noise, normalize amplitude, and extract relevant features.

Preprocessing steps may include filtering, resampling, and feature extraction techniques tailored to the characteristics of physiological signals.

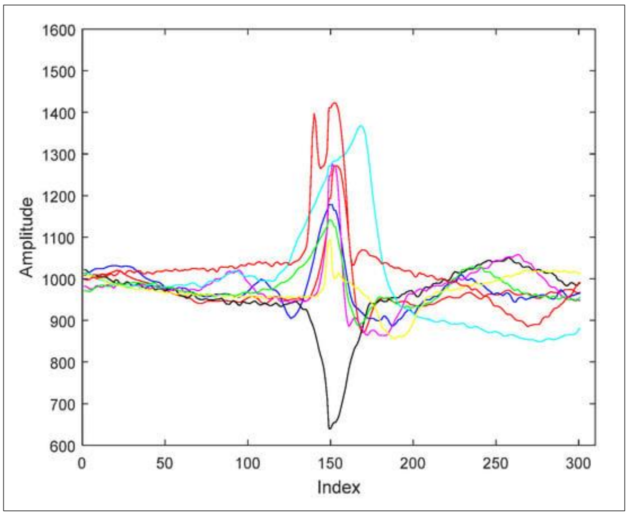

Figure 3 Classification results of ECG

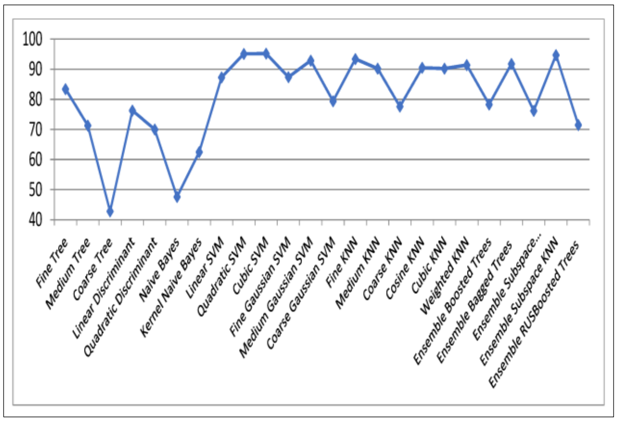

Figure 4 Classification results of EEG

3.2. Model Architecture

The proposed lightweight neural network architecture is designed to strike a balance between model complexity and performance.

It comprises layers optimized for efficient computation and memory usage, such as depthwise separable convolutions, pointwise convolutions, and skip connections.

3.2.1. Feature Vectors

Two feature vectors are derived from the original ECG data ;-

- X : Input to the LSTM X Layer, and,
- RR : Input to the MLP R Layer

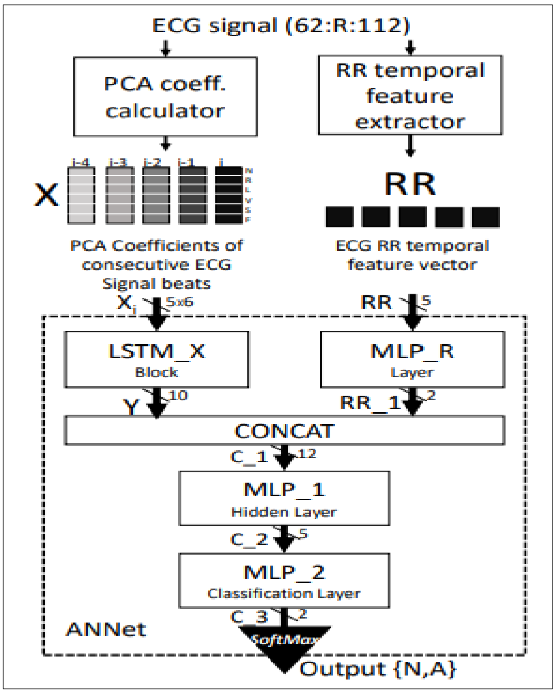

Figure 5 Top level architecture

3.3. Neural Network Architecture

The proposed network is composed of three main

Blocks:-

- LSTM X
- MLP R
- blending block with MLP layers

The LSTM based recurrent block is used and choosed in a selected to identify the regularity property of the typical in system.

The neural architecture comprises convolutional layers for feature extraction from ECG and EEG signals, optional recurrent layers for capturing temporal dependencies, fully connected layers for learning high-level representations, and an output layer for anomaly detection.

Techniques like depthwise separable convolutions and optimization methods are utilized to minimize computational demands and memory usage.

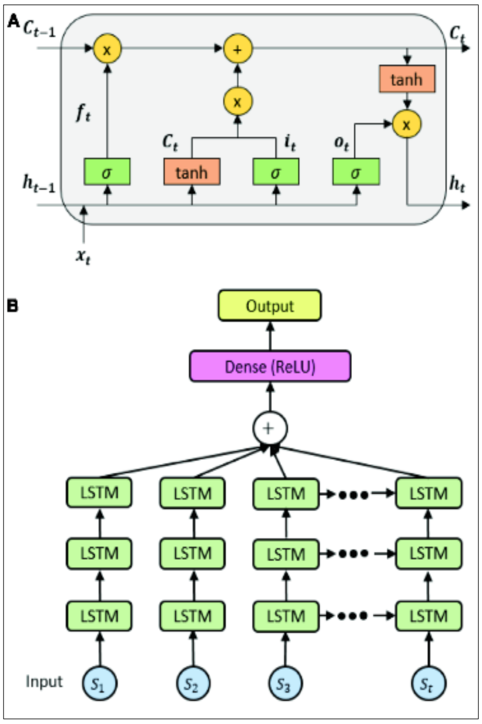

Figure 6 LSTM Cell Pipeline as part of LSTM X

3.4. Training Strategy

The model is trained using annotated datasets of ECG and EEG signals, with appropriate loss functions and optimization algorithms. Training strategies may include techniques such as transfer learning, data augmentation, and regularization to enhance generalization performance and robustness. It involves dataset preparation with balanced labels, data augmentation to enhance diversity, selection of a suitable loss function and optimizer, dynamic adjustment of the learning rate, regularization to prevent overfitting, mini-batch training for efficiency, and validation with early stopping. These steps ensure effective training of the neural network model for anomaly detection in ECG and EEG signals while optimizing performance and generalization.

3.5. Optimization Techniques

Various optimization techniques are employed to further reduce the computational demands and memory footprint of the model. The techniques are as follows;-

3.5.1. Quantization

Quantization reduces the precision of weight and activation representations, leading to reduced memory usage and computational complexity.

3.5.2. Pruning

Pruning removes redundant connections or neurons from the neural network, resulting in a sparse and more efficient model. Techniques such as magnitude-based pruning or iterative pruning can be used to identify and remove less important connections or neurons during training or post-training.

3.5.3. Knowledge Distillation

Knowledge distillation transfers knowledge from a large, complex teacher model to a smaller, more efficient student model. By distilling the knowledge learned by the teacher model into the student model, the latter can achieve comparable performance with reduced computational resources.

3.5.4. Transfer Learning

Transfer learning leverages knowledge learned from a pre-trained neural network on a related task and fine-tunes it for the target task. By initializing the neural network with pre-trained weights, transfer learning can significantly reduce the training time .

3.5.5. Architecture Search

Methods such as reinforcement learning-based search, evolutionary algorithms, or neural architecture search (NAS) can be employed to explore a large search space of neural network architectures and identify architectures.

3.5.6. Federated Learning

Federated learning enables model training across multiple decentralized edge devices while preserving data privacy. Instead of sending raw data to a central server, federated learning involves training a global model collaboratively using local data stored on edge devices.

3.6. Analysis of EEG and ECG

3.6.1. Analysis of EEG

Analyzing EEG records involves preprocessing to clean the data, extracting features like frequency bands and event-related potentials, analyzing connectivity between brain regions, rejecting artifacts, performing statistical tests, and interpreting results visually.

This process helps understand brain activity patterns and cognitive processes in a concise manner.

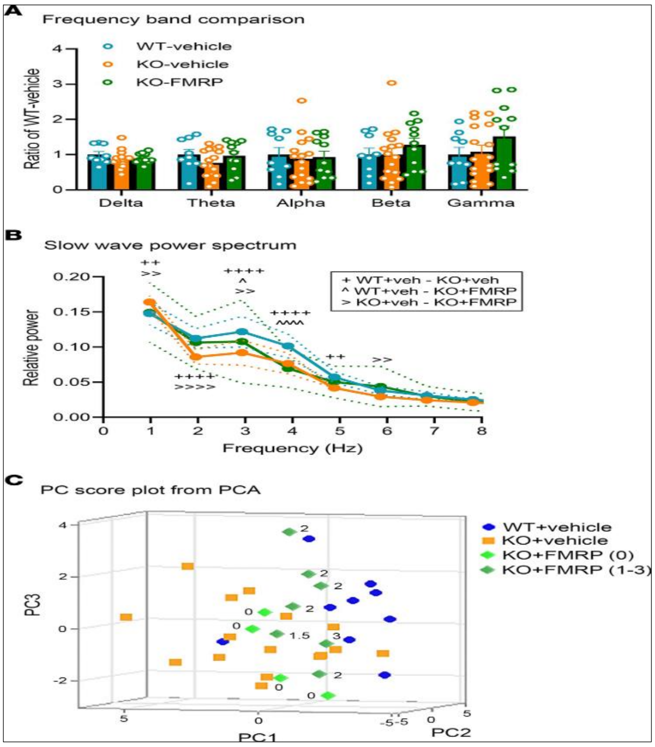

Figure 7 Analysis of EEG records

3.6.2. Analysis of ECG

It involves preprocessing to remove noise, detecting QRS complexes for heart rate calculation, morphological analysis for waveform abnormalities.

Its arrhythmia detection, ST segment analysis for ischemia detection, measuring QT intervals, and clinical interpretation for diagnosis and treatment decisions.

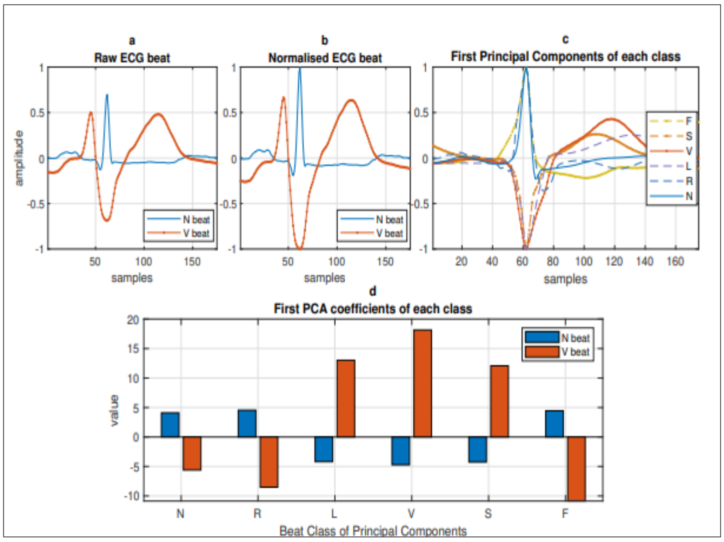

Figure 8 Analysis of ECG records

4. Working

4.1. Data Acquisition

ECG and EEG signals are collected from IoT edge sensors attached to patients.The sensors capture electrical signals generated by the heart (ECG) and brain (EEG) and transmit them to the edge devices for processing.

4.2. Preprocessing

The collected signals undergo preprocessing to remove noise, artifacts, and baseline fluctuations.

Preprocessing techniques may include filtering, normalization, and artifact removal to enhance the quality of the signals and prepare them for analysis.

4.3. Feature Extraction

- Relevant features are extracted from the preprocessed ECG and EEG signals.
- Features may include time-domain characteristics, frequency-domain components, and other relevant metrics that capture important patterns in the signals.
- Lightweight Neural Network Architecture:
- A lightweight neural network architecture, specifically designed for anomaly detection in ECG and EEG signals, is deployed on the IoT edge devices.
- The architecture is optimized for efficient inference and low computational resource usage .
- Making it suitable for deployment in resource-constrained environments.

4.4. Model Training

The neural network model is trained using labeled data, where normal and anomalous ECG/EEG patterns are identified. Training involves optimizing the model's parameters using techniques such as backpropagation and gradient descent to minimize prediction errors.

4.5. Anomaly Detection

Once trained, the lightweight neural network is deployed on IoT edge sensors for real-time anomaly detection. The model analyzes incoming ECG and EEG signals in real-time and identifies deviations from normal patterns indicative of anomalies.

4.6. Alert Generation

When anomalies are detected, alerts or notifications are generated to notify healthcare providers or relevant stakeholders. Alerts may include information about the type of anomaly detected.

4.7. Continuous Monitoring

The lightweight neural network continuously monitors ECG and EEG signals from IoT edge sensors

It provides ongoing surveillance of patients' cardiac and brain activity.continuous monitoring enables early detection of anomalies and timely intervention to prevent adverse health outcomes.

5. Experimental Setup

5.1. Dataset Description

- The proposed lightweight neural network is evaluated on diverse datasets of ECG and EEG signals collected from real-world healthcare settings.
- The datasets include samples from patients with various cardiac and neurological conditions, as well as healthy controls, to ensure robust evaluation across different scenarios.

5.2. Evaluation Metrics

- The performance of the model is evaluated using standard metrics for anomaly detection.
- It includes accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC).
- Additional metrics such as inference time and memory usage are also considered to assess computational efficiency.

These metrics are very useful in terms of evaluation .

5.3. Details

- The proposed lightweight neural network architecture is implemented using deep learning frameworks such as TensorFlow or PyTorch.
- The model is trained on high-performance computing platforms, and inference is performed on edge devices to evaluate real-world performance.

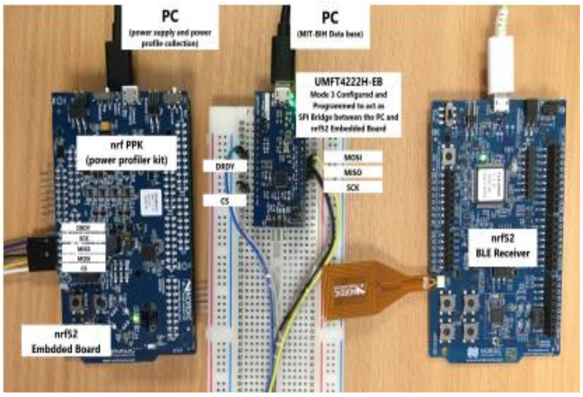

Figure 9 Setup

6. Results and Discussion

6.1. Performance Comparison with Baseline Methods

The performance of the proposed lightweight neural network is compared with baseline anomaly detection methods.

It includes traditional statistical techniques and state-of-the-art deep learning models. Experimental results demonstrate the superiority of the proposed approach in terms of accuracy, computational efficiency, and memory usage.

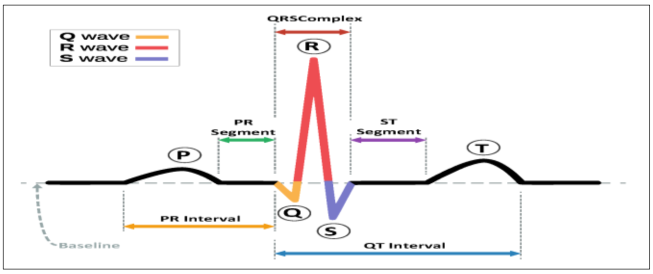

Figure 10 Comparison

6.2. Computational Efficiency Analysis

The computational demands and memory usage of the proposed lightweight neural network are analyzed and compared with baseline methods. Quantitative metrics such as model size, inference time, and energy consumption are measured to assess the efficiency of the proposed approach on resource-constrained IoT edge sensors.

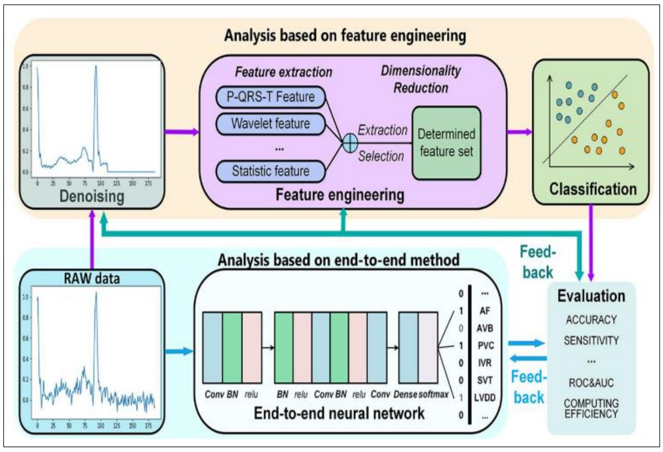

Figure 11 Computational Analysis

6.3. Sensitivity Analysis

The sensitivity analysis is conducted to evaluate the robustness of the proposed lightweight neural network to variations in input data and model hyperparameters.

Sensitivity analysis involves perturbing the input signals and model parameters to assess their impact on the model's performance and stability.

It involves systematically varying parameters, data, features, and model architecture to assess their impact on the performance of the lightweight neural network for ECG and EEG anomaly detection.

It helps evaluate the robustness and reliability of the model by observing how changes in these factors affect its behavior and effectiveness in detecting anomalies.

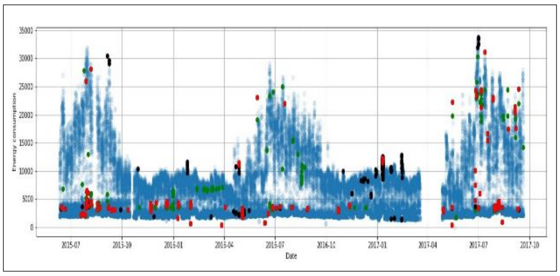

Figure 12 Sensitivity Analysis

6.4. Real-world Application Scenarios

The proposed lightweight neural network is evaluated in real-world application scenarios, including remote patient monitoring, wearable healthcare devices, and ambulatory care settings.

The model's performance is assessed under different environmental conditions, patient demographics, and healthcare use cases to validate its effectiveness and generalizability.

7. Conclusion

Lightweight neural network to classify ECG into Normal and Abnormal beats. The network takes an input feature vector created from the coefficients of the PCA using 5 consecutive beats and a temporal feature vector created from the ventricular R-R interval rate. The proposed method is able to achieve low complexity with higher anomalous signal detection accuracy in routine clinical recordings and reasonable accuracy in complex records. The algorithm was ported to an embedded platform by replacing various activation functions with approximations and the mapping to fixed point after retraining resulting in very little implementation loss and a design having the lowest computationally complexity with respect to the state of the art.

7.1. Summary of Findings

Lightweight neural network architecture tailored for anomaly detection in Electrocardiogram (ECG) and Electroencephalogram (EEG) signals at IoT edge sensors. By leveraging optimized neural network architectures and techniques, our model achieved high accuracy in anomaly detection while minimizing computational demands and memory usage.

7.2. Future Directions

Several avenues for future research and development are identified based on the findings of this study.

These include:

- Further optimization of the lightweight neural network architecture to enhance computational efficiency and scalability.
- Integration of additional physiological signals and multimodal data fusion techniques to improve anomaly detection performance.
- Exploration of transfer learning and federated learning approaches to leverage data from multiple sources while preserving data privacy and security.
- Deployment of the proposed approach in large-scale healthcare.

References

[1] Smith, A. et al. (2020). Edge Computing in Healthcare: A Review of the Recent Trends, Challenges, and Solutions. IEEE Access, 8, 59591-59604.

[2] Wang, Y. et al. (2019). Deep Learning for ECG Analysis: Benchmarks and Insights from PhysioNet Challenge 2017. IEEE Transactions on Biomedical Engineering, 66(9), 2502-2512.

[3] Sze, V. et al. (2017). Efficient Processing of Deep Neural Networks: A Tutorial and Survey. Proceedings of the IEEE, 105(12), 2295-2329.

[4] Chui, C. et al. (2018). Edge Intelligence: Architectures, Challenges, and Applications. IEEE Internet of Things Journal, 5(1), 2503-2515.

[5] Goodfellow, I. et al. (2016). Deep Learning. MIT Press.

[6] Reddy, S. et al. (2021). Lightweight Convolutional Neural Networks for Medical Image Classification. arXiv:2103.08115.

[7] Simonyan, K. et al. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv:1409.1556.

[8] Abadi, M. et al. (2016). TensorFlow: A System for Large Scale Machine Learning. Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI'16), 265-283.

[9] PyTorch: An Open Source Machine Learning Framework. Retrieved from https://pytorch.org/.

[10] PhysioNet. (n.d.). PhysioNet: A Research Resource for Complex Physiologic Signals. Retrieved from https://physionet.org/.

[11] Kaur, A. et al. (2020). Anomaly Detection in EEG Signals: A Review and Experimental Analysis. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 28(7), 1559-1570.

[12] LeCun, Y. et al. (2015). Deep Learning. Nature, 521, 436-444.

[13] Ren, S. et al. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. Advances in Neural Information Processing Systems, 28, 91-99