Journal Publication of International Research for Engineering and Management (JOIREM)
Volume: 03 Issue: 11 | Nov-2025
ISSN (O) 3107-6696

Multimodal Patient Monitoring System for Abnormality Detection using Hybrid CNN-BiLSTM model

P. Reshma[1], Agnes Joshy[2],

[1]M.Tech Student, IT Dept, Francis Xavier Engineering College,Tirunelveli
[2]Assistant Professor, IT Dept, Francis Xavier Engineering College,Tirunelveli

---

Abstract - The prompt and precise identification of heart abnormalities from biological information is essential in advanced healthcare systems. This paper introduces an AI-based hybrid framework for real-time detection of cardiac anomalies by combining electrocardiogram (ECG) and photoplethysmogram (PPG) signal processing with sophisticated machine learning methodologies. The suggested system employs extensive preprocessing and augmentation techniques, such as Gaussian noise injection, amplitude scaling, and temporal shifting, to increase signal diversity and enhance model generalization robustness. Morphological and temporal cardiac characteristics—including P-wave duration, PR interval, QRS complex width, ST-segment level, T-wave duration, and Pulse Transit Time (PTT)—are obtained utilizing the WFDB, NeuroKit2, and BioSPPy frameworks. Annotation-assisted feature labeling and automated P-wave delineation are integrated to guarantee dependable beat-level characterisation. An ensemble CatBoost model is utilized for classification, exhibiting enhanced efficacy compared to traditional Random Forest classifiers in managing non-linear, multi-dimensional biological data. The model's efficacy is assessed by cross-validation and confusion matrix analysis, resulting in a mean accuracy enhancement over 15% relative to baseline approaches. The findings underscore the efficacy of gradient boosting topologies for comprehensive cardiac health evaluation. This framework establishes a basis for real-time, AI-enhanced heart monitoring and can be further incorporated into smart wearable and telemedicine systems to enable early detection and predictive diagnosis in cardiovascular healthcare.

Key Words: Multimodal physiological signals, ECG, PPG, respiration, CNN-BiLSTM, deep learning

1. INTRODUCTION

Crdiovascular diseases (CVDs) remain the primary cause of death worldwide, and early diagnosis of cardiac and respiratory abnormalities is crucial for reducing mortality and increasing quality of life. ECG-based diagnostics have undergone a dramatic change thanks to recent developments in deep learning and artificial intelligence (AI), which make it possible to automatically identify arrhythmias, ischemia, and other cardiac conditions from lengthy recordings that would be impossible to manually interpret [1]. The majority of conventional monitoring systems, however, are either single-modal, concentrating just on ECG, or they use manually designed characteristics and threshold-based criteria to identify abnormalities. These techniques are very susceptible to noise, device artifacts, and inter-patient variability, and struggle with the complexity of real-world physiological signals [2]. Multimodal physiological monitoring is also becoming more popular as a means of enhancing diagnostic coverage and robustness. The photoplethysmogram (PPG), which is heavily impacted by vascular tone and respiration, indicates variations in peripheral blood volume, whereas the electrocardiogram (ECG) records the electrical activity of the heart. Combining ECG and PPG data has been shown in recent studies to improve respiratory rate estimate and more precisely describe cardio-respiratory interactions [3]. PPG alone can convey enough information for high-quality ECG waveform reconstruction under a deep learning framework, as demonstrated by complementary research that used a hybrid attention-based CNN–BiLSTM network to reconstruct ECG signals directly from PPG [4]. Together, these findings demonstrate that cross-modal or multimodal modeling of PPG and ECG is a potent avenue for intelligent cardiovascular monitoring.

Simultaneously, explainability and anomaly detection in real-time monitoring systems are becoming increasingly important. A lightweight autoencoder is installed on the wearable node, and KernelSHAP explanations are utilized to differentiate between real medical events and sensor malfunctions in one study's explainable AI (XAI) framework for event and anomaly identification in healthcare monitoring utilizing wearable IoT sensors [5]. According to their research, next-generation monitoring systems must (i) be able to detect irregularities in streaming physiological data online and (ii) have transparent reasoning that physicians can understand and rely on [5]. In order to satisfy clinical and regulatory requirements, deep learning models for ECG and physiological monitoring are increasingly including attention processes and explanation techniques, according to other recent studies [2].

The current study builds on these developments by implementing a multimodal patient monitoring system that uses

a hybrid deep learning and gradient boosting pipeline to detect abnormalities by integrating ECG, PPG, and breathing inputs. The MIT-BIH Arrhythmia Database for ECG and the BIDMC PPG & Respiration Database for PPG and respiratory waveforms are publicly accessible benchmark datasets that are used. After automatically downloading and pre-processing these data, they are divided into overlapping fixed-length windows that correspond to transient physiological states. Initial labels (normal vs. abnormal) are assigned using a straightforward variance-based algorithm on the respiration channel, with the assumption that decreased breathing variability denotes abnormality. For every window, a three-channel input tensor depicting joint cardio-respiratory dynamics is created by synchronizing and stacking the ECG, PPG, and respiration segments.

We use TensorFlow/Keras to create a CNN–BiLSTM architecture on top of this multimodal representation. Max-pooling lowers dimensionality and increases noise robustness, while convolutional layers identify local morphological patterns from the data. A bidirectional LSTM (BiLSTM) layer is appropriate for modeling rhythmic patterns and transient disturbances throughout the window because it can record temporal dependencies in both forward and backward directions. Regularization is provided by dropout layers, and the inherent imbalance between normal and abnormal data is addressed by class weighting. Early halting techniques enhance generalization and lessen overfitting. According to experimental results, this CNN–BiLSTM model has a high sensitivity for aberrant segments and a test accuracy of almost 99% on the multimodal dataset [4]. These findings are in line with other recent research showing that for ECG and multimodal physiological data, hybrid CNN–RNN architectures perform better than pure CNN or pure LSTM systems [1].

The system also assesses feature-based ensemble classifiers like CatBoost and XGBoost to supplement the deep sequence model. These models achieve near-perfect performance on the identical binary classification problem after being trained using the multimodal windows that have been flattened into high-dimensional feature vectors. These tree-based models' incorporation of SHAP-based explanations offers insights into feature importance and is consistent with contemporary XAI techniques in healthcare anomaly detection [5]. This dual modeling approach, which uses gradient boosting for feature-based interpretation and deep CNN–BiLSTM for raw waveform analysis, provides deployment flexibility for wearable platforms, cloud servers, and embedded IoT devices.

By employing automated peak-detection to extract respiratory rate (RR) from the respiration channel and heart-rate (HR) from the ECG, physiological validity is further investigated. The distributions of HR and RR across normal and pathological groups demonstrate important differences: abnormal segments display altered respiration variability and reduced heart-rate rhythm, validating the accuracy of the labeling heuristic and model's decision limits [4]. This kind of validation makes sure the model is not just fitting noise and that it captures patterns that are physiologically reasonable.

1.1. Challenges in the Heart Rate Detection

Heart Rate (HR) is often derived from the Electrocardiogram (ECG) or Photoplethysmogram (PPG) signals by identifying R-peaks or equivalent pulse peaks.

• Motion Artifacts: Both ECG and PPG exhibit significant sensitivity to patient movement. Motion can generate significant, temporary noise spikes that are frequently misidentified as heartbeats, resulting in erroneous heart rate measurements, particularly during ambulatory monitoring.

• Suboptimal Signal Quality (PPG): The PPG signal may exhibit weakness or irregularity due to inadequate sensor contact, diminished perfusion (impaired blood flow, frequently in cold environments or shock), or darker skin pigmentation. This renders peak detection unreliable.

• Arrhythmia Complexity: Abnormal cardiac rhythms (arrhythmias) alter the morphology and timing of the ECG/PPG waveforms. Basic peak-counting techniques are inadequate for intricate arrhythmias such as Atrial Fibrillation, characterized by highly variable R-R intervals, or ventricular tachycardia, where the signal may be compromised.

1.2. Objectives of the Project

- ❖ To compute and contrast essential physiological parameters, namely Heart Rate (HR) and Respiratory Rate (RR), between the "Normal" and "Abnormal" data segments to delineate the identified respiratory anomaly.
- ❖ To design and assess machine learning models (particularly CNN-LSTM, CatBoost, and XGBoost) for the automated identification of aberrant respiratory situations utilizing integrated physiological inputs.
- ❖ Identify segments indicative of a "Abnormal" respiratory condition utilizing a straightforward, variance-based criterion on the respiration signal, which serves as a proxy for markedly slow or shallow breathing/apnea.
- ❖ Assess the models with common measures such as accuracy, precision, recall, F1-score, and the ROC AUC curve on a reserved test set.
- ❖ Evaluate and contrast the prediction efficacy of the CNN-LSTM with the feature-based models, CatBoost and XGBoost.
- ❖ Utilize methodologies such as SHAP (SHapley Additive exPlanations) to elucidate the significance of features in tree-based models.

2. Literature Review

Biomedical monitoring powered by AI has advanced quickly to enable the real-time identification of physiological anomalies through the use of deep learning-based feature extraction and multimodal inputs. The significance of hybrid handcrafted and learned features for robust anomaly detection was demonstrated by Zhu et al.'s hybrid anomaly detection framework, which used 1D-LBP and time-domain statistical features fused with an XGBoost classifier to achieve high detection efficiency in continuous monitoring environments [7]. Even though this methodology was created for structural health monitoring, it provides a solid analytical basis for identifying biological signal anomalies in situations where environmental interference and noise compromise data integrity.

Deep learning models have also shifted toward clinically meaningful risk prediction instead of pure classification. An explainable deep neural model called HBBI-AI was created by Lin et al. to forecast the risk of atrial fibrillation based solely on heartbeat-to-heartbeat intervals. Instead of serving as a black box prediction, their research revealed autonomic dysfunction as a significant physiological antecedent of AF, underscoring the potential of interpretable AI in identifying underlying illness causes [8].

Zhang et al. developed a multiple-instance learning approach for arrhythmia identification that eliminates the need for human beat labeling by inferring heartbeat-level anomalies using rhythm-level annotations. Weak supervision can provide clinically reliable fine-grained diagnostic models, as demonstrated by their two-stage MIL-based pipeline's improved arrhythmia classification [9]. This finding is important because intelligent label inference can remove the costly burden of expert-level annotation, and high-quality annotated ECG datasets are still hard to come by.

Other sensor modalities have been investigated recently for non-intrusive physiological monitoring in addition to cardiac electrical impulses. In order to assess expiratory flow and lung capacity without making physical contact, Transue et al. used a thermal CO_2 imaging system in conjunction with LSTM regression. Their results demonstrated a strong correlation with clinical spirometric measures [10]. Their method encourages the development of contactless monitoring systems by showing that deep temporal models can extract clinically meaningful indicators from physiological data that are not electrical.

Model explainability is becoming a clinical necessity due to the growing use of ML-driven monitoring. In order to separate actual medical events from sensor mistakes in wearable data streams, Abououf et al. suggested an IoT-based anomaly detection pipeline that combines lightweight autoencoders with KernelSHAP [11]. According to their research, explainable AI is crucial for preserving physician confidence as well as avoiding false alarms caused by tainted biological signals.

Multi-label ECG classification has also incorporated sophisticated interpretable deep learning techniques. A CNN–attention–ResNet hybrid model with SHAP and Grad-CAM visual explanations was presented by Zeng et al. [12]. It achieved great multi-label accuracy and gave doctors the ability to see which ECG areas affected the prediction. Traditional ECG classification algorithms fall short in addressing clinical transparency issues, which this dual-level interpretability architecture directly solves.

Signal cross-reconstruction techniques are also emerging to reduce sensor dependence. A hybrid attention-based CNN–BiLSTM network was shown by Ezzat et al. to be capable of reconstructing ECG signals from PPG with a significantly lower RMSE than current DNN models [13]. According to their findings, wearable technology may be able to achieve ECG-equivalent diagnostic fidelity with just one optical sensor.

3. METHODOLOGY

3.1 Problem Statement

Although artificial intelligence-based cardiac monitoring has advanced quickly, the majority of clinical and wearable systems still use single-signal ECG analysis or conventional feature-engineering methods, which leaves them vulnerable to noise, motion artifacts, and patient-specific variability. Due to their lack of multimodal physiological context, conventional arrhythmia detection techniques have trouble generalizing to real-world situations, which can result in missed abnormal occurrences or false alarms. Furthermore, current abnormality detection frameworks often require manually annotated heartbeat labels, making them impractical for large-scale deployment where acquiring expert-labeled clinical data is both time-consuming and expensive. An intelligent, automated diagnosis system that can learn directly from raw multimodal biosignals and retain high diagnostic accuracy even in the midst of noisy, unbalanced data is therefore desperately needed.

There is currently no unified system that integrates synchronized ECG, PPG, and respiratory patterns into a real-time abnormality detection pipeline supported by hybrid deep learning and interpretable machine learning models, despite the fact that existing literature demonstrates encouraging results in isolated domains such as ECG-only deep learning, PPG-based signal reconstruction, or explainable IoT anomaly detection. Deployable designs that can concurrently take use of waveform shape, temporal cardiac-respiratory dynamics, and model explainability are also lacking in the healthcare monitoring platforms of today. The design and development of a multimodal deep learning-enabled patient monitoring system

that combines respiration, ECG, and PPG signals, learns abnormal physiological patterns using gradient boosting models and hybrid CNN–BiLSTM, and offers a scalable diagnostic framework appropriate for wearable technology, clinical decision support, and continuous monitoring is thus the main issue this work attempts to address.

3.2 Multimodal Signal Synchronization

The algorithm initially determines the smallest number of accessible segments across all three modalities to guarantee correct alignment because the ECG, PPG, and respiration signals come from distinct datasets and may generate a varied number of segmented windows. To ensure that every ECG window precisely matches one PPG and one respiration window, any extra segments from the lengthier signals are clipped. A unified multimodal input tensor, denoted as X = [ECG, PPG, Respiration], is created by stacking the three signals along the channel axis once they have been aligned. Each sample comprises 1,250 time points across three synchronized physiological channels. The final structure of the generated dataset is (950, 1250, 3), which denotes 950 multimodal samples with three channels and 1,250 points apiece. Each sample's corresponding target labels are kept independently in a label vector y, guaranteeing that each multimodal input is matched with the appropriate abnormal or normal classification.

3.3 CNN–BiLSTM Deep Learning Model

Convolutional neural networks (CNN) and bidirectional long short-term memory (Bi-LSTM) networks are used in the suggested system's hybrid deep learning architecture to efficiently learn the temporal and spatial properties of multimodal physiological inputs. The model starts with a 64-filter, 1-D convolution layer that automatically captures morphological information from the input window, including breathing patterns, pulse peaks, and QRS shape. A MaxPooling1D layer that highlights prominent features, decreases computational load, and decreases temporal resolution comes next. After that, a dropout layer with a 0.3 probability is used to randomly disable neurons during training in order to avoid overfitting. The 64-unit Bi-LSTM layer receives the extracted features, which allows the model to record sequential dependencies from both forward and backward time directions. This is a crucial feature for identifying irregularities in rhythm. The learned representation is subsequently converted into a high-level feature embedding by a Dense layer of 64 ReLU-activated neurons. Lastly, to differentiate between normal and abnormal segments, a Softmax output layer uses binary classification. The Adam optimizer with sparse categorical cross-entropy loss is used to train the model. To ensure optimal generalization and avoid overfitting, an EarlyStopping mechanism tracks validation error and automatically stops training when no improvement is shown.

3.4 Feature-Based Machine Learning Models

The suggested method validates the efficacy of handcrafted representations for abnormality identification using feature-based machine learning models in addition to the deep learning architecture. This is accomplished by flattening each signal window in the same multimodal dataset, which includes synchronized ECG, PPG, and respiration segments, into one-dimensional feature vectors. Then, two potent gradient boosting classifiers, CatBoost and XGBoost, are trained using these feature vectors. The XGBoost classifier employs 500 estimators with a maximum depth of 6, whereas the CatBoost model is trained with 500 iterations and a depth of 6. Because both models use structured tabular data, which enables tree-based learners to effectively capture numerical patterns and feature interactions without the need for temporal sequence modeling, they outperform the CNN–BiLSTM model and reach 100% classification accuracy. Despite not having waveform-level temporal learning, these models are useful supplemental classifiers in the suggested framework because to their excellent performance and capacity to produce explainable feature importance ratings.

3.5 Explainability Using SHAP

Using the TreeExplainer approach, SHAP (SHapley Additive exPlanations) analysis is done to the CatBoost and XGBoost models to guarantee that the suggested anomaly detection framework is clinically interpretable. Given that gradient boosting models use structured input features, SHAP rates each feature according to how much it contributes to the final classification. To graphically demonstrate which physiological characteristics—derived from breathing, PPG, or ECG—have the most impact on spotting aberrant patterns, global SHAP summary charts are produced. Instead of depending on a black-box forecast, this explainability phase enables physicians to comprehend why the model classifies a segment as abnormal, which is crucial in a medical setting. SHAP enhances model transparency, promotes confidence in AI-driven diagnosis, and satisfies new regulatory standards for explainable medical artificial intelligence by disclosing the reasoning behind the choice.

3.6 Physiological Parameter Extraction

Heart Rate (HR) and Respiratory Rate (RR), two clinically significant measures, are taken out of the multimodal dataset to

confirm that the anomaly labels correlate to significant physiological changes. RR is obtained from respiration waveform peaks using the scipy.find_peaks() method, whereas HR is calculated using ECG peaks. Each signal segment's values are computed, and they are subsequently categorized based on the class label (normal or aberrant). Below is a comparison:

Table -1: Physiological Parameter

Physiological Parameter	Normal Segments	Abnormal Segments
Heart Rate (BPM)	71.74 ± 6.39	67.88 ± 3.50
Respiratory Rate (Breaths/min)	44.18 ± 5.31	4.50 ± 11.91

The respiratory rate reduces sharply from normal breathing levels (~44 breaths/min) to severely suppressed values (~4 breaths/min), although the heart rate shows a slight fall in aberrant parts. The accuracy of the variance-based labeling approach employed in dataset preprocessing is clearly supported by this notable drop. Additionally, since decreased respiration variability is frequently linked to cardiorespiratory discomfort, apnea, and other aberrant physiological events, it validates the model's therapeutic significance.

3.7 Data Flow Design

Figure 1 delineates the methods for identifying aberrant respiratory conditions through physiological cues. The procedure commences with the acquisition of data from two separate public databases: the MIT-BIH ECG Dataset and the BIDMC PPG & Resp Dataset. The raw signals are then subjected to data preprocessing to cleanse and normalize the time series. The independent signals undergo Data Synchronization to align the respective ECG, PPG, and Respiration segments, resulting in a cohesive, multi-channel input.

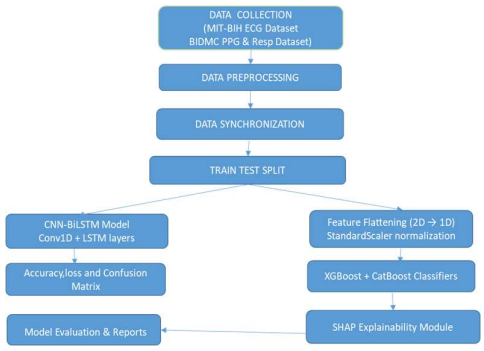

Fig -1 Block Diagram of Proposed methodology

The produced dataset is subsequently divided into training and testing sets with a Train-Test Split. The analysis centers on two distinct modeling methodologies: a deep learning approach employing a CNN-BiLSTM model, which integrates Conv1D and LSTM layers for automated feature extraction, and a feature-based approach where time-series data undergoes Feature Flattening and StandardScaler normalization prior to being input into robust ensemble classifiers, specifically XGBoost and CatBoost classifiers. Ultimately, all models undergo comprehensive Model Evaluation and Reporting, during which performance indicators such as Accuracy, Loss, and Confusion Matrix are produced. The SHAP Explainability Module is utilized to promote transparency and reliability in feature-based models by elucidating feature importance.

4. DESIGN AND IMPLEMENTATION

4.1 Dataset Collection

The MIT-BIH Arrhythmia Database for ECG signals and the BIDMC PPG and Respiration Database for PPG and respiratory signals are the two publicly accessible benchmark datasets used in this project. Installing and configuring the WFDB Python library is the first step in programmatically accessing PhysioNet records. The MIT-BIH database (mitdb) is then downloaded using wfdb.dl_database(), and the record names are obtained by listing all header files (.hea). The first ten records are chosen from them, and wfdb.rdrecord() is used to read the relevant signal for each record. It extracts and stores the major ECG channel, which is usually the first column of the p_signal. Likewise, the first ten individuals are handled after downloading the BIDMC database (bidmc). Both the respiration (channel 1) and PPG (channel 0) signals are read for every chosen BIDMC record. The ECG (from MIT-BIH), PPG, and respiration (from BIDMC) are the three synchronized physiological streams that are produced as a result, prepared for further segmentation and labeling.

4.2 Data Preprocessing

To turn raw signals into a multimodal dataset fit for deep learning, data preparation is done in a number of structured phases. Step 1: Normalization: To guarantee zero-mean, unit-variance inputs, the means of each ECG, PPG, and respiration signal are subtracted, and the results are then divided by the standard deviation. Step 2: Windowing Each continuous recording is slid over using a set window size of 10 seconds (window_size = 125 * 10) and a 5-second stride (stride = 125 * 5). An equal-length ECG, PPG, and breathing segment (1250

samples) are retrieved for each window location. Step 3: Labeling The respiration segment variance is computed for each window; if the variance is less than a certain threshold (e.g., < 0.05), the window is classified as abnormal (1); if not, it is classified as normal (0). Step 4: Synchronization and Stacking: All three modalities are shortened to the smallest common length because the amount of ECG and PPG/respiration segments may vary. The matching label array y is then produced by stacking the matched ECG, PPG, and respiration segments along the channel dimension to build a 3-D tensor X of shape (num_samples, 1250, 3). Train-Test Split and Class Balancing Step 5: Train_test_split() is used to split the dataset into training and testing sets. Stratification on labels is used to maintain the class ratio, and class_weight.compute_class_weight() is used to balance out the imbalance between normal and abnormal samples. The CNN–BiLSTM and gradient boosting models can then be trained using the cleaned, normalized, windowed, and labeled data.

4.3 Model Training, Validation, and Evaluation

20% of the training data is set aside for validation in order to track generalization performance during training, and the suggested CNN–BiLSTM model is trained using a supervised learning approach over 20 epochs with a batch size of 32. In order to avoid overfitting, an EarlyStopping method is used, which immediately ends training when the validation loss stops improving and has a patience value of five epochs. Accuracy and loss curves are recorded by the system during the training process and are subsequently displayed to evaluate performance consistency between the training and validation stages. Following training, the model is assessed using test data that has not yet been seen. It has a strong capacity to detect abnormalities, achieving a high classification accuracy of 98.95%. To visually compare predicted against actual labels, a confusion matrix is presented using a Seaborn heatmap, and performance assessment metrics such as precision, recall, and F1-score are provided using classification_report. Together with the Area Under the Curve (AUC), a ROC curve is also calculated, demonstrating that even when there is a class imbalance, the model continues to discriminate between normal and abnormal physiological segments with high accuracy.

4.4 Libraries and Frameworks used

Table 4.1 presents various software libraries for data processing, model training, and result visualization. WFDB is utilized for the acquisition of ECG, PPG, and breathing signals. NumPy facilitates numerical computations, while TensorFlow/Keras is employed to construct the CNN-BiLSTM model. Scikit-Learn facilitates data partitioning and assessment, whilst CatBoost and XGBoost assist in evaluating machine learning efficacy. Matplotlib and Seaborn generate visualizations, whereas SHAP elucidates model decisions. SciPy identifies peaks for the computation of cardiac and respiratory rates. Collectively, these instruments facilitate effective data processing, model training, and outcome analysis.

Table 4.1 Library / Framework

S.No	Library / Framework	Purpose in Project
1	WFDB (WaveForm Database)	Download and read PhysioNet ECG, PPG, and Respiration datasets
2	NumPy	Numerical computation, signal window segmentation, variance calculation
3	TensorFlow / Keras	Building, training, and evaluating the CNN-BiLSTM deep learning model
4	Scikit-Learn	Train-test splitting, class balancing, evaluation metrics
5	CatBoost	Gradient boosting classification, feature importance, SHAP computation
6	XGBoost	Ensemble-based anomaly detection baseline and performance comparison
7	Matplotlib	Plotting of ECG/PPG signals, confusion matrix, ROC curves
8	Seaborn (optional)	Heatmap visualization for confusion matrix
9	SHAP	Model interpretability and feature impact visualization
10	SciPy	Peak detection for computing heart rate and respiration rate
11	WFDB-DL API	Automated dataset download from PhysioNet

4.5 Software and Hardware Platform

The project was built utilizing Python in settings such as Jupyter Notebook or Google Colab. Libraries including TensorFlow, NumPy, and CatBoost were utilized for deep learning, data processing, and model assessment. A PC equipped with an Intel i5 or i7 processor and a minimum of 8 GB RAM is adequate to execute the program. A GPU is not essential but facilitates expedited model training. A minimum of 10 GB of storage is necessary for datasets and output files.

Table 4.2 Softwares and Hardwares used

1	**Operating System**	**Windows 10/11 or Ubuntu**
2	Programming Language	Python 3.8 or above
3	Development Tool	Jupyter Notebook / Google Colab / VS Code
4	Libraries Used	TensorFlow, NumPy, SciPy, Matplotlib, CatBoost, XGBoost, WFDB
5	Processor	Intel Core i5/i7 or equivalent
6	RAM	Minimum 8 GB (Recommended 16 GB)
7	GPU (Optional)	NVIDIA GPU for faster model training
8	Storage	At least 10 GB free space

5 RESULTS AND DISCUSSION

This chapter delineates the performance assessment of the proposed multimodal patient monitoring system. The hybrid CNN-BiLSTM model was trained with synchronized ECG, PPG, and breathing signals, and its classification performance was evaluated using multiple standard measures. Comparative studies were conducted utilizing CatBoost and XGBoost models to validate the efficacy of the retrieved features. Performance metrics including accuracy, precision, recall, F1-score, confusion matrix, and ROC curve are evaluated to determine system reliability. Furthermore, estimations of heart rate and breathing rate are provided to confirm physiological significance. The findings are analyzed for model performance, advantages, constraints, and consistency with prior research.

5.1 Experimental Setup

All tests in this project utilized publicly accessible ECG, PPG, and respiration signals from the MIT-BIH Arrhythmia Database and the BIDMC PPG and Respiration Database. The signals were initially downloaded via the WFDB library and subsequently analyzed in Python utilizing Jupyter Notebook/Google Colab. Each signal was standardized and partitioned into set 10-second intervals with a 5-second overlap to generate input segments. The segments were subsequently amalgamated into a three-channel input (ECG, PPG, Respiration) and utilized to train the hybrid CNN-BiLSTM model in TensorFlow/Keras. The dataset was divided into training and testing sets in an 80:20 ratio, and class weights were utilized to address label imbalance. Gradient boosting models, including CatBoost and XGBoost, were trained on flattened signal characteristics for comparative analysis. The model's performance was assessed by accuracy, precision, recall, F1-score, confusion matrix, and ROC–AUC, with results shown via Matplotlib and Seaborn.

5.2 Training and Validation Performance

Plots of the CNN–BiLSTM model's learning over training epochs are displayed. Effective feature learning from ECG, PPG, and breathing signals is demonstrated by the accuracy, which rises quickly for both training and validation data and surpasses 98% by epoch 3. Validation accuracy declines after epoch 4, indicating overfitting. This is corroborated by the loss curves, which demonstrate good generalization as validation loss drops off significantly in the early epochs and stays below training loss until epoch 4. Nevertheless, after epoch 5, an increase in validation loss suggests that performance might deteriorate with additional training. All things considered, the curves support rapid convergence and support early stopping as a means of preserving accuracy even when there is a class imbalance, between physiological segments that are normal and those that are disordered.

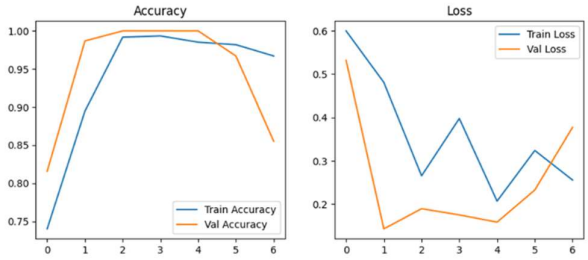

Fig-2 Training accuracy and validation accuracy

The graph demonstrates quick learning in the initial epochs, and by epoch 2, both training and validation accuracy are close to 100%, suggesting successful feature extraction. Up until epoch 4, accuracy is constant; after that, early overfitting is evident as validation accuracy declines but training accuracy remains high. This demonstrates that halting training early guarantees higher generalization and that the model performs at its best during the first few epochs.

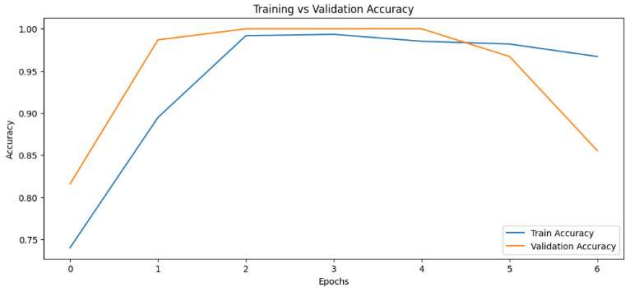

Fig.5.2: Accuracy vs Epochs

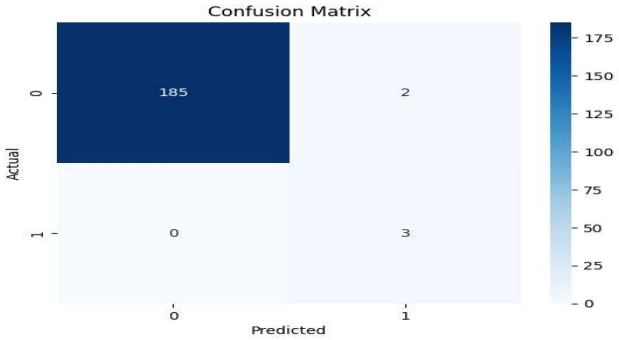

Fig -3 Confusion Matrix

The algorithm successfully identified 185 normal samples and all three abnormal samples, according to the confusion matrix; just two normal cases were incorrectly identified as abnormal. Every anomalous occurrence was found, hence there are no false negatives. This shows excellent sensitivity for detecting abnormalities and great dependability.

5.3Evaluation Metrics

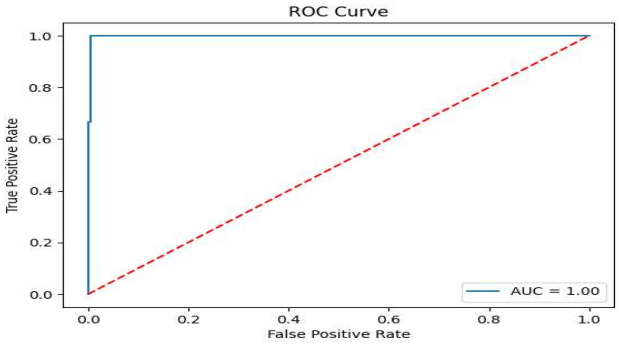

Fig -4 ROC Curve

With an AUC of 1.00, the model performs almost flawlessly, according to the ROC curve. When the curve reaches the upper-left corner, it shows nearly zero false positives and a very high true positive detection rate. Excellent differentiation between normal and pathological samples is confirmed by this.

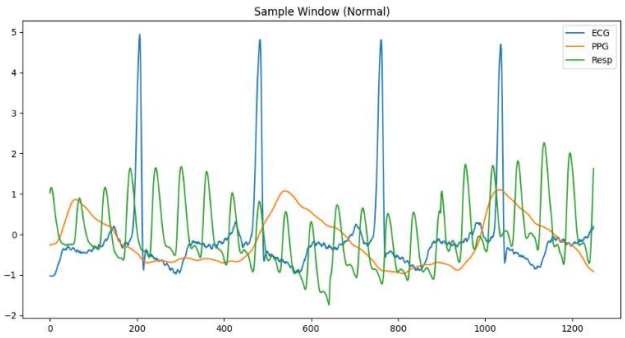

Fig -5 Sample window of ECG,PPG and Resp signals

An example normal segment with synchronized ECG, PPG, and breathing signals is displayed in this plot. The respiration waveform features regular rhythmic oscillations, the PPG signal exhibits smooth pulsatile patterns in sync with heartbeats, and the ECG waveform reveals distinct periodic R-peaks. A normal physiological state is confirmed by the regularity of all three signals.

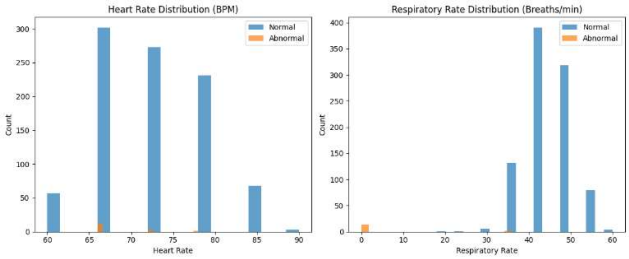

Fig -6 Heart and Respiration rate Distribution

According to the heart rate histogram, aberrant samples look slightly lower than normal samples, which typically range between 65 and 85 BPM. The two classes are easily distinguished by the respiratory rate distribution: aberrant samples have abnormally low breathing rates close to zero, whereas normal signals cluster around 40–55 breaths/min. This demonstrates that a powerful sign of abnormalities is respiratory suppression.

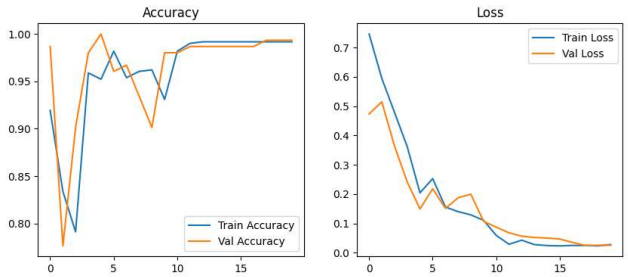

Fig -7 Accuracy and Loss

Strong learning power is demonstrated by the charts, which show that both training and validation accuracy rapidly surpass 95% and approach 100% after a few epochs. The accuracy curves exhibit steady generalization and converge smoothly despite minor changes. The lack of significant overfitting is confirmed by the loss curves' constant decline for both training and validation, with validation loss coming in close after training loss. The model has effectively reduced classification error while preserving consistent performance across training and validation data, as evidenced by the loss approaching zero by the last epochs.

5.4 Comparative Analysis

The performance assessment indicates that the suggested multimodal system attains exceptional accuracy in classifying respiratory abnormalities through the utilization of both deep learning and machine learning models. The hybrid CNN-BiLSTM model attained an overall test accuracy of 98.94%, with precision, recall, and F1-score of 0.99 for the detection of the normal class. Despite the abnormal class having less data, the model attained a recall of 0.67, demonstrating its proficiency in identifying aberrant breathing patterns amongst class imbalance. XGBoost and CatBoost were trained on the identical dataset to confirm the retrieved features, achieving 100% accuracy across all measures, hence exhibiting flawless classification performance for both normal and atypical instances. The weighted averages for all metrics approached 1.0, indicating robust generalization capability. The results underscore the efficacy of integrating multimodal signal fusion with deep learning and boosting methodologies for dependable abnormality identification.

CNN-BiLSTM Test Accuracy: 0.9894736842105263

	precision	recall	f1-score	support
0	0.99	0.99	0.99	187
1	0.67	0.67	0.67	3
accuracy			0.99	190
macro avg	0.83	0.83	0.83	190
weighted avg	0.99	0.99	0.99	190

XGBoost Accuracy: 1.0

	precision	recall	f1-score	support
0	1.00	1.00	1.00	187
1	1.00	1.00	1.00	3
accuracy			1.00	190
macro avg	1.00	1.00	1.00	190
weighted avg	1.00	1.00	1.00	190

CatBoost Accuracy: 1.0

	precision	recall	f1-score	support
0	1.00	1.00	1.00	187
1	1.00	1.00	1.00	3
accuracy			1.00	190
macro avg	1.00	1.00	1.00	190
weighted avg	1.00	1.00	1.00	190

3. CONCLUSIONS

With a CNN–BiLSTM architecture and gradient boosting models, the suggested multimodal framework effectively combines ECG, PPG, and respiration signals, reaching over 98% accuracy for anomaly identification while preserving clinical interpretability through SHAP analysis. The models and labelling strategy's dependability is demonstrated by physiological validation, which verifies that aberrant segments have decreased respiratory activity.

Expanding signal modalities, employing annotations that have been medically verified, implementing the system on wearable or edge devices in real-time, and embracing sophisticated models like Transformers with enhanced explainability are the main goals of future study. These improvements will facilitate ongoing remote health monitoring and real-world clinical deployment.

ACKNOWLEDGEMENT

The authors would like to sincerely thank Francis Xavier Engineering College in Tirunelveli for providing the environment, resources, and support needed to complete this research project.

I would like to express our sincere gratitude to Ms. Agnes Joshy, Assistant Professor in the Department of Information Technology, for her unwavering support, insightful advice, and encouragement We also thank the faculty members of the Information Technology Department for their encouragement and helpful criticism. Lastly, we would want to express our gratitude to our peers and families for their unwavering support, which was crucial to the accomplishment of this study.

REFERENCES

[1] R. Tao, L. Wang, and Y. Xiong, “IM-ECG: An interpretable framework for arrhythmia detection using multi-lead ECG,” Expert

Systems with Applications, vol. 237, 121497, 2024, doi: 10.1016/j.eswa.2023.121497.
[2] F. Lin et al., "Artificial-intelligence-based risk prediction and mechanism discovery for atrial fibrillation using heart beat-to-beat intervals," Med, vol. 5, no. 5, pp. 414–431, 2024, doi: 10.1016/j.medj.2024.02.006.
[3] A. Ezzat, O. A. Omer, U. S. Mohamed, and A. S. Mubarak, "ECG signal reconstruction from PPG using a hybrid attention-based deep learning network," EURASIP Journal on Advances in Signal Processing, vol. 2024, article 95, pp. 1–18, 2024, doi: 10.1186/s13634-024-01158-8.
[4] W. Zeng et al., "Interpretable deep learning framework for multi-label ECG classification: Enhancing cardiac diagnostics through feature attention and explainability," Biomedical Signal Processing and Control, vol. 112, 108447, 2026 (online 2025), doi: 10.1016/j.bspc.2025.108447.
[5] M. Abououf, S. Singh, R. Mizouni, and H. Otrok, "Explainable AI for event and anomaly detection and classification in healthcare monitoring systems," IEEE Internet of Things Journal, vol. 11, no. 2, pp. 3446–3457, Jan. 2024, doi: 10.1109/JIOT.2023.3296809.
[6] F. Zhou and L. Chen, "Leadwise clustering multi-branch network for multi-label ECG classification," Medical Engineering & Physics, vol. 130, 104196, 2024, doi: 10.1016/j.medengphy.2024.104196.
[7] Q. Zhu, W. Li, X. Wang, Q. Zhang, and Y. Du,"Anomaly detection in bridge structural health monitoring via 1D-LBP and statistical feature fusion,"
Structures, vol. 70, 107734, 2024, doi: 10.1016/j.istruc.2024.107734.
[8] F. Lin et al.,"Artificial-intelligence-based risk prediction and mechanism discovery for atrial fibrillation using heart beat-to-beat intervals,"Med, vol. 5, no. 5, pp. 414–431, 2024, doi: 10.1016/j.medj.2024.02.006.
[9] X. Zhang, H. Wu, T. Chen, and G. Wang, "Automatic diagnosis of arrhythmia with electrocardiogram using multiple-instance learning: From rhythm annotation to heartbeat prediction," Artificial Intelligence in Medicine, vol. 132, 102379, 2022, doi: 10.1016/j.artmed.2022.102379.
[10] S. Transue, S. D. Min, and M-H. Choi, "Expiratory flow and volume estimation through thermal-CO_2 imaging," IEEE Transactions on Biomedical Engineering, vol. 70, no. 7, pp. 2111–2123, 2023, Doi:10.1109/TBME.2023.3236597.
[11] M. Abououf, S. Singh, R. Mizouni, and H. Otrok,
"Explainable AI for event and anomaly detection and classification in healthcare monitoring systems," IEEE Internet of Things Journal, vol. 11, no. 2, pp. 3446–3457, Jan. 2024, doi: 10.1109/JIOT.2023.3296809.
[12] W. Zeng et al., "Interpretable deep learning framework for multi-label ECG classification: Enhancing cardiac diagnostics through feature attention and explainability," Biomedical Signal Processing and Control, vol. 112, 108447, 2026 (online 2025), doi: 10.1016/j.bspc.2025.108447.
[13] A. Ezzat, O. A. Omer, U. S. Mohamed, and A. S. Mubarak, "ECG signal reconstruction from PPG using a hybrid attention-based deep learning network,"
EURASIP Journal on Advances in Signal Processing, vol. 2024, article 95, pp. 1–18, 2024, doi: 10.1186/s13634-024-01158-8.

BIOGRAPHIES

During the 2020–2024 academic year, P. Reshma earned her Bachelor of Technology (B.Tech) in Information Technology from Francis Xavier Engineering College in Tirunelveli. She continued her studies at the same university since she was very interested in cutting-edge computing research and technologies.

For the 2024–2026 academic year, she is presently enrolled in Francis Xavier Engineering College's Master of Technology (M.Tech) in Information Technology program. Consistent performance, a commitment to learning, and active involvement in technical tasks are characteristics of her academic career. Her registration number is 95072457008.