

A Hybrid Multi-Modal Deep Learning Architecture for Early Detection of Post-COVID Cardiac Disorders

D. Jeyabharathi ^{1,*}, N. N. Krishna Veni ², E. Sucila³

^{1,*} Research Scholar (Reg.No:20122052282006), Department of Computer Science - Mathematics, G. Venkataswamy Naidu College, Kovilpatti, 628 502, Affiliated to Manonmaniam Sundaranar University, Abishekapatti - 627 012, Tirunelveli, Tamil Nadu, India.

² Assistant Professor, Department of Computer Science, Holy Cross Home Science College, Thoothukudi, 626140, Tirunelveli, Tamil Nadu, India. Affiliated to Manonmaniam Sundaranar University, Abishekapatti, 627 012, Tirunelveli, Tamil Nadu, India

³ Associative Professor, Department of Mathematics, G. Venkataswamy Naidu College, Kovilpatti, 628 502, Affiliated to Manonmaniam Sundaranar University, Abishekapatti, 627 012, Tirunelveli, Tamil Nadu, India.³

Received: 25.01.2024

Revised: 30.02.2024

Accepted: 15.03.2024

Abstract

The COVID-19 pandemic has resulted in a rise in post-acute cardiac cases such as myocarditis and arrhythmias, thus emphasizing the need for early detection systems. The traditional methods of diagnosis are not sensitive enough to detect subclinical conditions, thus causing delays in treatment and higher morbidity rates. To overcome the above challenges, this research work proposes a hybrid multi-modal deep learning architecture for the early and accurate detection of post-COVID cardiac disorders. This research work proposes a hybrid deep learning architecture (CNN-LSTM) for the early detection of post-COVID cardiovascular disorders. An ensemble module (CapsNet, DenseNet-121, NASNet) is used to obtain multi-level feature sets, thus improving the robustness of predictions. An adaptive score-based integration technique is used to combine the predictions, thus improving the accuracy of classification and preventing misclassifications. The proposed framework is tested for its efficacy in performing early and accurate cardiovascular risk detection. The experimental outcome shows that the proposed architecture performs better with an accuracy of 91.3%, precision of 90.9%, recall of 89.8%, and F1-score of 94.2%, which is better than the existing models. The enhanced diagnostic results show the effectiveness of the proposed hybrid multi-modal model for the accurate diagnosis of early post-COVID cardiovascular abnormalities. The study demonstrates the potential of hybrid multi-modal deep learning to transform post-COVID cardiac surveillance, offering scalable, early, and reliable diagnostic support for preventive cardiology and risk stratification in clinical practice.

Keywords: Post-COVID cardiovascular disorder, Hybrid deep learning, CNN-LSTM, Ensemble learning, Adaptive score integration, Early disease detection.

1. Introduction

Coronavirus Disease 2019 (COVID-19) has resulted in severe short-term and long-term health issues worldwide, with cardiovascular diseases being identified as one of the most severe post-recovery complications [1]. Recent clinical studies have shown that patients who have recovered from COVID-19 are at a higher risk of developing cardiac abnormalities such as myocarditis, arrhythmias, heart failure, and thromboembolic events [2]. These complications have been linked to inflammatory reactions, immune system imbalances, and endothelial damage caused by the SARS-CoV-2 virus [3]. Early diagnosis of cardiovascular diseases post-COVID-19 is therefore crucial in minimizing mortality rates and ensuring improved patient outcomes. However, traditional diagnosis methods are highly dependent on clinical assessments, imaging, and biomarker analysis, which may not be able to detect the slightest changes in cardiovascular diseases, hence the need for the incorporation of advanced computational techniques [4].

Recently, artificial intelligence (AI) and deep learning algorithms have received considerable attention for the automated diagnosis of diseases and analysis of medical data [5]. Convolutional Neural Networks (CNNs) have been widely used for the diagnosis of cardiovascular diseases because of their capability to extract spatial features from medical images and physiological signals such as electrocardiograms (ECGs) and echocardiographic information [6], [7]. Various research studies have shown that CNN models provide better diagnostic results for the identification of cardiac anomalies. Although CNN models are highly effective, their architectures mainly concentrate on the extraction of spatial features and lack the capability to extract temporal features from patient health records, which are necessary for tracking the progression of post-COVID cardiovascular diseases [8].

The Convolutional Neural Networks (CNNs) have been successful in the extraction of spatial features from cardiac imaging modalities, whereas the Recurrent Neural Networks (RNNs) and their variants, including Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, have been efficient in the modeling of temporal dependencies in sequential data such as ECG signals [9]. However, the existing applications of AI have largely been restricted to single-modality analysis. Cinar et al. [10] proposed a CNN to detect left ventricular dysfunction in sinus rhythm ECGs, whereas Sowmya et al. [11] proposed a 3D CNN for automatic cardiac MRI segmentation. Although effective, these single-modality analyses have not been able to extract the holistic, multi-dimensional pathophysiological pattern of post-COVID cardiac injury, which is manifested across imaging, electrophysiological, and clinical modalities [12].

Adaptive fusion techniques have been introduced recently to overcome the drawbacks of traditional ensemble techniques. Score-level fusion techniques dynamically weight the predictions of models according to confidence levels and performance criteria, thus enhancing decision-making accuracy and minimizing misclassification errors [13]. Although the hybrid deep learning approach, multi-modal integration, and adaptive ensemble techniques have shown promising

outcomes, very few studies have explored the fusion of these techniques for early detection of post-COVID cardiovascular diseases.

Overall, the growing availability of various types of clinical data, such as electronic health records, ECG signals, cardiac imaging, and biomarker data, presents substantial opportunities for the early detection of post-COVID cardiac irregularities [10-13]. However, the majority of the existing models are developed to work with single sources of data and are not capable of handling multiple sources of clinical data with different formats and characteristics. This leads to the existing models being prone to decreased prediction accuracy, poor generalization performance, and clinical reliability [14]. Thus, there is a pressing need to design a hybrid multi-modal deep learning model that can handle multiple sources of clinical data to accurately and reliably detect early post-COVID cardiovascular complications.

To address these research gaps, this paper presents a Hybrid Multi-Modal Deep Learning Architecture for early detection of post-COVID cardiac disorders. The proposed system combines hybrid CNN-based sequential learning models with ensemble deep learning models and adaptive score-level fusion techniques to improve classification accuracy and robustness. The proposed system is designed to develop a trustworthy and efficient diagnostic tool for early cardiovascular risk assessment in post-COVID patients. Figure 1 represents graphical view of hybrid multi-modal deep learning architecture for early cardiac disease Detection

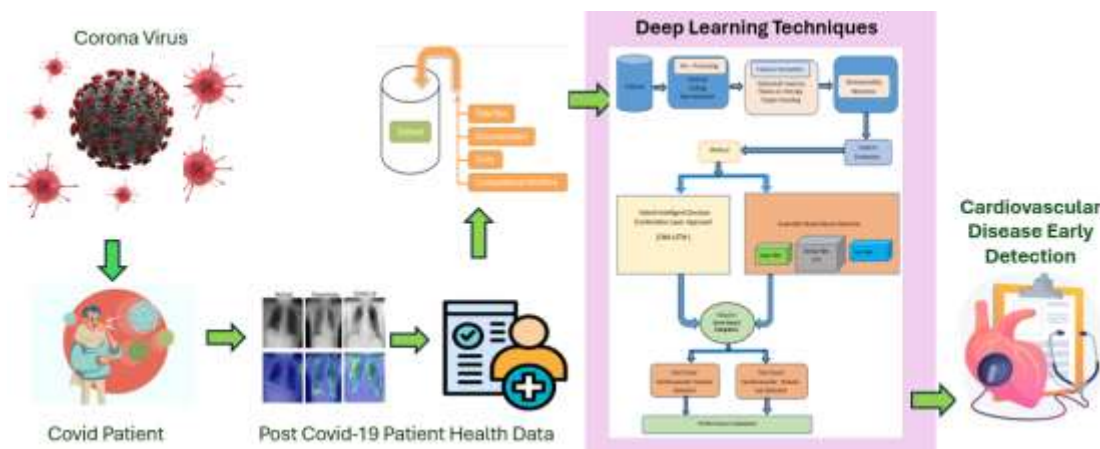


Figure 1: Graphical schematic diagram of deep learning architecture based early cardiac disease detection

1.1 Novelty and Contributions of this Study:

- Designed a hybrid multi-modal deep learning model that combined CNN, and LSTM to extract spatial and temporal information from the post-COVID clinical data.

10.48047/jocaaa.2024.33.02.45

- Incorporated the latest ensemble models (CapsNet, DenseNet-121, and NASNet) to enhance the ability of feature learning and the robustness of classification.
- Proposed an adaptive score-level fusion approach to assign dynamic weights to the model outputs, which can improve the accuracy of predictions and prevent misclassifications.

2. Related Works

Recent research has shown that post-COVID patients are increasingly experiencing cardiovascular complications such as myocarditis, arrhythmia, heart failure, and thrombotic disorders due to inflammation and immune responses caused by the virus [1-4]. To validate early diagnosis, artificial intelligence and deep learning algorithms have been extensively used for the detection of cardiac diseases [8-10]. CNNs have been successfully used for spatial feature extraction from ECG signals and medical images, while RNNs, including LSTM and GRU networks, have the capability to process temporal and sequential patient data [11,12]. More recently, transformer models and ensemble learning models have been proposed to enhance the accuracy of predictions by identifying complex relationships between features and aggregating the predictions of multiple models [9]. Multi-modal learning models have also received attention for their ability to combine clinical data, imaging, laboratory values, and physiological signals for a more comprehensive analysis of patient health [13]. Fusion models, including feature-level fusion, decision-level fusion, and score-level fusion, have also demonstrated clinical benefits by enhancing the accuracy of diagnosis and preventing misclassification [14]. However, the existing literature may use single-modal datasets, static fusion approaches, or simple hybrid models, which emphasize the requirement for sophisticated multi-modal deep learning models with adaptive fusion techniques for effective early diagnosis of post-COVID cardiovascular diseases [15].

3. Methodology

3.1 Proposed Hybrid Multi-Modal Deep Learning Architecture

The proposed research work suggests the use of a Hybrid Multi-Model Deep Learning Architecture for the early diagnosis of post-COVID cardiovascular diseases. The proposed architecture involves a series of stages, such as data preprocessing, feature extraction, dimensionality reduction, hybrid model training, ensemble-based detection, decision-level fusion, and performance analysis. Figure 2 represents the proposed hybrid multi-modal deep learning architecture

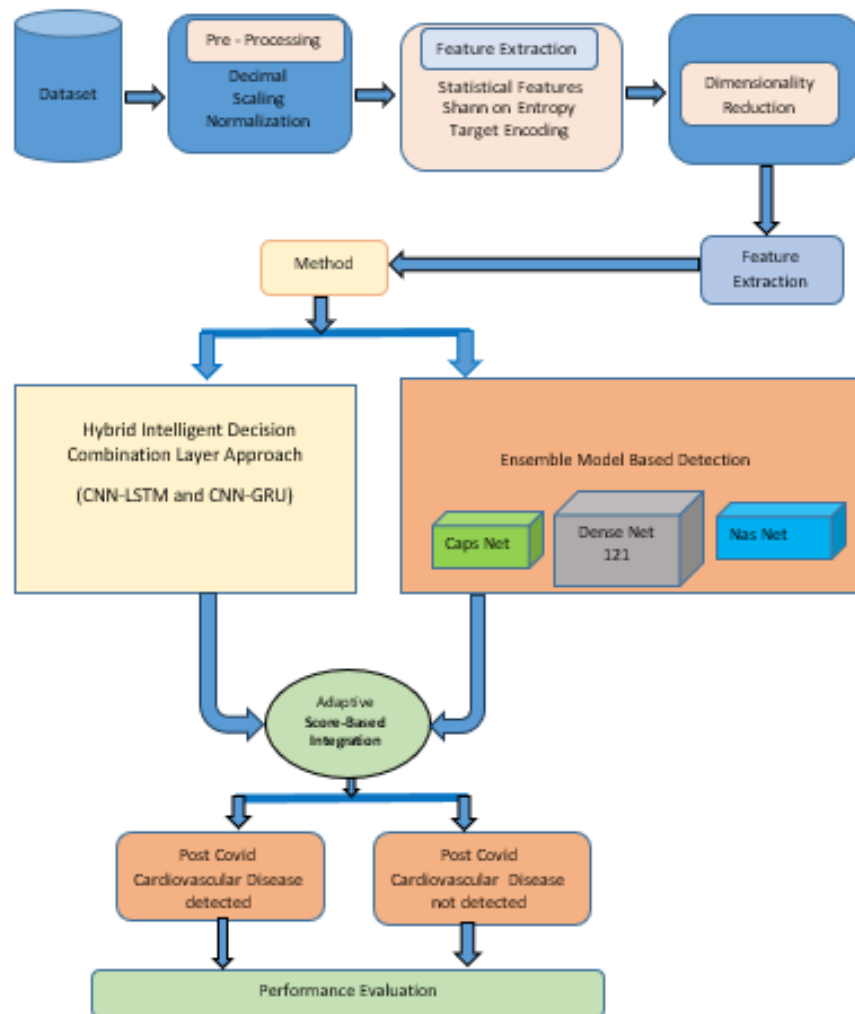


Figure 2: Proposed Hybrid multi-modal deep learning architecture.

3.1. Data Pre-Processing

The acquired dataset is first processed through a data preprocessing step to improve data quality and training efficiency of the proposed architecture [16,17]. Missing and inconsistent data are handled using appropriate imputation methods. Normalization and scaling processes are further carried out to normalize the feature space and remove data bias. Table 1 represents the demographic and health profile of study participants.

Table 1: Demographic and Health Profile of Study Participants

Gender	Age Group	Education	Heart Disease	Diabetes	Other Disease	Smoking	Blood Pressure	Weight	Vaccinated	Vaccination Dose Status	Anxiety Before	Anxiety After
Female	10–20	Higher Study	Yes	Yes	Yes	Yes	Low	High	Yes	1 & 2 Dose	Neutral	Neutral
Female	41–50	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Agree	Agree
Male	10–20	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1 & 2 Dose	Agree	Agree
Female	31–40	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Agree	Strongly Agree
Female	21–30	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Neutral	Neutral
Female	31–40	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Neutral	Agree
Male	31–40	Higher Study	Yes	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Disagree	Agree
Male	21–30	Higher Study	No	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Disagree	Disagree
Female	31–40	Higher Study	No	Yes	Yes	Never	Low	High	Yes	1, 2 & Booster	Agree	Strongly Agree
Male	21–30	Higher Study	No	Yes	Yes	Yes	Low	High	Yes	1, 2 & Booster	Neutral	Disagree

The preprocessing step helps in improving the training convergence speed and preventing the proposed architecture from overfitting due to irregular variations in the acquired data.

Noise Removal

To eliminate high-frequency artifacts from medical imaging data while preserving critical cardiac structural information, we applied anisotropic diffusion filtering [1], which smooths homogeneous regions while preserving edges through the partial differential equation:

$$\frac{\partial I}{\partial t} = \nabla \cdot (c(\|\nabla I\|)\nabla I) \quad (1)$$

where I represents the image intensity, t is the diffusion time, and $c(\cdot)$ is the edge-stopping function that prevents smoothing across boundaries.

Missing Data Handling

10.48047/jocaaa.2024.33.02.45

For incomplete clinical parameters (e.g., missing troponin values, omitted ECG leads), we employed Multiple Imputation by Chained Equations (MICE) [2], generating $m = 5$ complete datasets through iterative regression models:

$$X_{\text{miss}}^{(t+1)} = f(X_{\text{obs}}, X_{\text{imp}}^{(t)}; \theta) \quad (2)$$

where missing values X_{miss} are sequentially imputed based on observed data X_{obs} and previously imputed values $X_{\text{imp}}^{(t)}$, with parameters θ estimated from the complete cases.

Data Normalization

Multi-modal feature scales were harmonized using modality-specific normalization: medical images underwent Z-score normalization per channel ($\mu = 0, \sigma = 1$), while clinical variables were min-max scaled to $[0,1]$:

$$\hat{x}_i = \frac{x_i - \min(\mathbf{X})}{\max(\mathbf{X}) - \min(\mathbf{X})}, \text{ for clinical data} \quad (3)$$

$$\hat{I}(x, y) = \frac{I(x, y) - \mu_I}{\sigma_I}, \text{ for imaging data} \quad (4)$$

ensuring stable gradient flow during backpropagation across heterogeneous inputs.

Data Augmentation

To enhance model generalization and simulate clinical variability, we applied stochastic geometric and intensity transformations with probability $p = 0.5$, including random rotations ($\theta \sim \mathcal{U}(-15^\circ, 15^\circ)$), scaling ($s \sim \mathcal{U}(0.9, 1.1)$), and additive Gaussian noise ($\eta \sim \mathcal{N}(0, 0.01\sigma^2)$):

$$I_{\text{aug}} = T_{\text{geom}}(I) + \eta, \eta \sim \mathcal{N}(0, \sigma_{\text{noise}}^2) \quad (5)$$

where T_{geom} represents the composite geometric transformation, effectively expanding our training distribution to improve robustness against real-world acquisition variances.

3.2. Feature Extraction and Encoding

Feature extraction is carried out in order to extract meaningful clinical and physiological information. Statistical analysis is used to find important patterns and variability in patient health parameters. In addition, target encoding is used to represent categorical medical features as numerical values while retaining their predictive importance.

Statistical Feature Analysis

The study extracted 42 statistical features per modality, including central moments and dispersion measures. For a clinical parameter vector $\mathbf{x}$ of length N , we computed:

Skewness:

$$\gamma_1 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^3}{\sigma^3} \quad (6)$$

Kurtosis:

$$\beta_2 = \frac{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^4}{\sigma^4} - 3 \quad (7)$$

Coefficient of Variation:

$$CV = \frac{\sigma}{\mu} \times 100\% \quad (8)$$

where μ and σ are the mean and standard deviation of $\mathbf{x}$, respectively.

Target Encoding for Categorical Medical Variables

Let C be a categorical variable with K categories $\{c_1, c_2, \dots, c_K\}$ and target y (cardiac disorder: 0/1). The target-encoded value v_k for category c_k is a smoothed combination:

$$v_k = \lambda \cdot \frac{\sum_{i=1}^N \mathbb{I}(C_i = c_k) \cdot y_i}{\sum_{i=1}^N \mathbb{I}(C_i = c_k)} + (1 - \lambda) \cdot \bar{y} \quad (9)$$

where $\mathbb{I}(\cdot)$ is the indicator function, $\bar{y}$ is the global target mean, and $\lambda \in [0, 1]$ is a smoothing parameter based on category frequency:

$$\lambda = \frac{n_k}{n_k + \alpha} \quad (10)$$

with n_k as the count of category c_k , and $\alpha = 10$ as a regularization constant to prevent overfitting for rare categories.

Multi-Modal Feature Fusion Concatenation

$$\mathbf{F}_{\text{fused}} = \mathbf{S}_1 \cdot \mathbf{F}_{\text{image}} \oplus \mathbf{S}_2 \cdot \mathbf{F}_{\text{clinical}} \oplus \mathbf{S}_3 \cdot \mathbf{F}_{\text{signal}} \quad (11)$$

where $\oplus$ denotes concatenation, and $\mathbf{S}_i$ are learnable score weight matrices for each modality.

3.3. Dimensionality Reduction

After feature extraction, dimensionality reduction is carried out to eliminate redundant and irrelevant features. This step helps to reduce complexity and retain important diagnostic

10.48047/jocaaa.2024.33.02.45

information. The resulting feature space enhances the classification accuracy and generalization ability of the proposed deep learning architectures. Although the CNN autoencoder greatly reduces the dimensionality of the image, the deep features may still be redundant and correlated. To further optimize the deep feature space, Principal Component Analysis (PCA) is employed. If the CNN autoencoder's encoder first compresses the input image I to a latent vector h , and PCA further reduces it:

$$h = f_{\text{enc}}(I) \text{ (Autoencoder encoding)} \quad (12)$$

$$z = P \cdot (h - \mu_h) \quad (13)$$

where:

- f_{enc} is the CNN encoder function
- $P \in \mathbb{R}^{k \times m}$ is the PCA projection matrix ($k < m$, $m = \dim(h)$)
- μ_h is the mean of the latent vectors
- z is the final reduced feature vector for classification

PCA is used to transform the deep feature vectors into a lower-dimensional orthogonal space by retaining the principal components that capture the maximum variance. This removes irrelevant features, reduces complexity, and improves the generalization capability of the deep learning model. The PCA-reduced feature set is then fed into the classification network for precise prediction of post-COVID heart disease.

3.4. Hybrid Intelligent Decision Combination Layer

The optimized set of features is then processed through a hybrid deep learning layer that is intended to identify both spatial and temporal relationships in the clinical data. This layer combines CNN-LSTM models. CNN is responsible for identifying spatial features, while the long short-term memory (LSTM) networks are used to identify temporal relationships that exist in the patient health records. The hybrid combination helps to improve the identification of disease patterns and increases the accuracy of the model.

CNN Spatial Feature Extraction

Let the input clinical image or signal at time t be $X_t \in \mathbb{R}^{H \times W \times C}$, where H, W, C are height, width, and channels. The CNN applies a convolution operation to extract spatial features:

$$F_t^{(l)} = \sigma \left(W^{(l)} * F_t^{(l-1)} + b^{(l)} \right) \quad (14)$$

where:

- $F_t^{(0)} = X_t$
- $W^{(l)}$ and $b^{(l)}$ are the filter weights and bias at layer l
- $*$ denotes the convolution operation
- σ is the ReLU activation function
- $F_t^{(L)}$ is the final spatial feature map at time t after L layers

LSTM Temporal Sequence Modeling

The flattened spatial feature vector $f_t = \text{Flatten}(F_t^{(L)})$ is fed into the LSTM unit to capture temporal dependencies:

$$\begin{aligned}
 i_t &= \sigma(W_{xi}f_t + W_{hi}h_{t-1} + b_i) \\
 f_t^{(\text{gate})} &= \sigma(W_{xf}f_t + W_{hf}h_{t-1} + b_f) \\
 o_t &= \sigma(W_{xo}f_t + W_{ho}h_{t-1} + b_o) \\
 \tilde{c}_t &= \tanh(W_{xc}f_t + W_{hc}h_{t-1} + b_c) \\
 c_t &= f_t^{(\text{gate})} \odot c_{t-1} + i_t \odot \tilde{c}_t \\
 h_t &= o_t \odot \tanh(c_t)
 \end{aligned} \tag{15}$$

where:

- $i_t, f_t^{(\text{gate})}, o_t$ are input, forget, and output gates
- c_t is the cell state
- h_t is the hidden state (temporal feature representation)
- $\odot$ denotes element-wise multiplication

Hybrid CNN-LSTM Output Integration

The final decision for patient p across T time steps is obtained by aggregating the LSTM hidden states and applying a classification layer:

$$y_p = \sigma \left(W_y \cdot \frac{1}{T} \sum_{t=1}^T h_t + b_y \right) \tag{16}$$

where:

10.48047/jocaaa.2024.33.02.45

- h_t is the LSTM hidden state at time t (fed by CNN features f_t)
- W_y and b_y are weights and bias of the final dense layer
- σ is the sigmoid activation for binary classification
- y_p is the predicted probability of post-COVID cardiac disorder

3.5. Ensemble Model-Based Detection

To increase the accuracy of the classification model, an ensemble detection model is introduced. The ensemble model combines several advanced deep learning models, such as the Capsule Network (CapsNet), DenseNet-121, and NASNet models. The CapsNet model helps to improve hierarchical feature learning, while the DenseNet-121 model helps to improve feature propagation and reuse. The NASNet model helps to improve feature extraction through architecture search optimization. The ensemble model helps to learn from multi-level feature representations.

Score based Weighted Ensemble Voting

Let the predictions of the three base models be:

- **Capsule Network:** $P_C(x)$
- **DenseNet-121:** $P_D(x)$
- **NASNet:** $P_N(x)$

The final ensemble prediction $\hat{y}$ is given by:

$$\hat{y} = \sum_{m \in \{C,D,N\}} w_m \cdot P_m(x) \quad (17)$$

where:

- $w_C + w_D + w_N = 1$ (normalized weights)
- w_m is the weight for model m , learned during validation
- $P_m(x) \in [0,1]$ is the probability output of model m

Average Ensemble:

$$\hat{y} = \frac{P_C(x) + P_D(x) + P_N(x)}{3} \quad (18)$$

3.6. Adaptive Score-Based Integration

10.48047/jocaaa.2024.33.02.45

The results obtained from the hybrid and ensemble learning modules are integrated using an adaptive score-based integration approach. This approach automatically assigns weights to the predictions of individual models based on their confidence levels and performance contributions. The decision-making integration approach improve classification accuracy and minimizes misclassifications, ensuring the accurate identification of post-COVID cardiovascular abnormalities.

Let:

- s_H be the prediction score vector from the Hybrid CNN-LSTM model
- s_E be the prediction score vector from the Ensemble model (CapsNet + DenseNet + NASNet)
- α and β be adaptive weights based on model confidence and performance

The final integrated prediction score is:

$$\mathbf{s}_{\text{final}} = \alpha \cdot \mathbf{s}_H + \beta \cdot \mathbf{s}_E \quad (19)$$

where:

$$\alpha = \frac{\text{Confidence}_H}{\text{Confidence}_H + \text{Confidence}_E}, \beta = 1 - \alpha \quad (20)$$

and the confidence for each model is computed as:

$$\text{Confidence}_m = \exp\left(-\frac{\text{Error}_m}{\text{Error}_H + \text{Error}_g}\right) \quad (21)$$

with Error_m being the validation error (e.g., cross-entropy loss) of model m .

Final Decision Rule:

The predicted class is then:

$$\hat{y} = \begin{cases} 1 & \text{if } \mathbf{s}_{\text{final}} \geq 0.5 \\ 0 & \text{otherwise} \end{cases} \quad (22)$$

where 1 indicates post-COVID cardiovascular disease detected and 0 indicates not detected.

3.7. Classification and Performance Evaluation

The final integrated system classifies patient status into two classes: post-COVID cardiovascular disorder detected and post-COVID cardiovascular disorder not detected. The performance of the

10.48047/jocaaa.2024.33.02.45

proposed model is assessed using statistical parameters like accuracy, precision, recall, sensitivity, specificity, F1-score, false positive rate, and false negative rate. The performance parameters confirm the efficacy and correctness of the proposed Hybrid Deep SKBKA-CR model for the early detection of cardiovascular disorders.

4. Results and Discussion

4.1 System and Hardware and software configuration:

The proposed work was conducted using a high-performance computing system that is equipped with an Intel/AMD multi-core processor, 16-32 GB RAM, and GPU support for deep learning model training. The proposed hybrid multi-modal architecture was developed using the Python programming language. Deep learning frameworks such as TensorFlow and PyTorch were used for model development and analysis. Data preprocessing, visualization, and analysis were performed.

4.2 Evaluation of Metrics and Performance Analysis:

The performance evaluation of the proposed hybrid multi-modal deep learning model demonstrates superior results compared to conventional approaches. Table 2 represents the performance comparison of the proposed model.

Accuracy: The proportion of total correct predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (23)$$

Precision: The proportion of positive predictions that are correct.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (24)$$

Recall : The proportion of actual positives that are correctly identified.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (25)$$

F1-Score: The harmonic mean of Precision and Recall.

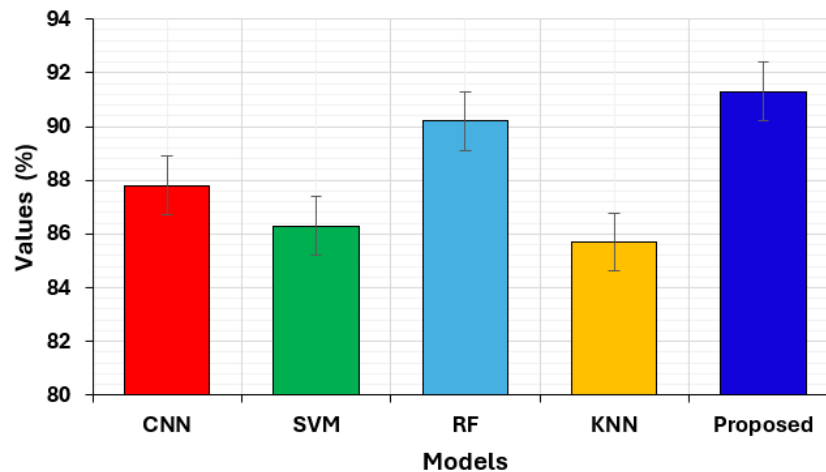
$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (26)$$

Table 2: Performance Comparison of the Proposed model

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
CNN	87.8	90.0	87.2	90.4
SVM	86.3	82.5	81.5	85.2
RF	90.2	83.7	85.6	86.7
KNN	85.7	81.9	80.6	83.1
Proposed	91.3	90.9	89.8	94.2

Accuracy :

The hybrid multi-modal deep learning model for the early detection of post-COVID cardiac disorders, the proposed model showed better classification performance with an accuracy of 91.3%. Figure 3 represents the accuracy analysis of the proposed model.

**Figure 3:** Accuracy of the Proposed Model

This is due to the efficient fusion of different clinical and physiological modalities. The model improves the recognition of patterns in complicated cardiac measures. Therefore, it provides more reliable results for early diagnosis than traditional models.

Precision (Prediction Reliability Perspective):

In terms of prediction reliability, the proposed model showed a precision of 90.9%, validating its effectiveness in correctly detecting actual cardiac disorder instances. The fusion of the multi-modal features decreases the chances of false positives in clinical screening. This provides more reliable positive predictions for post-COVID cardiovascular risk analysis. The finding validates the model’s effectiveness in facilitating accurate clinical decision-making. Figure 4 represents the precision analysis of the proposed model.

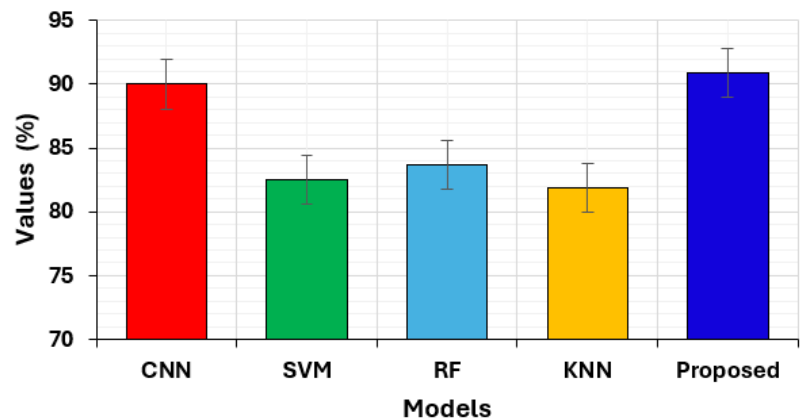


Figure 4: Precision the Proposed Model

Recall (Disease Detection Capability Perspective):

The proposed model had a recall of 89.8%, which indicates its high capability of detecting most of the actual post-COVID cardiac abnormal cases. This high detection rate indicates the efficiency of the architecture in reducing the number of missed disease cases. By using multi-source patient information, the system improves the detection of early-stage diseases. Figure 5 represents the recall analysis of the proposed model.

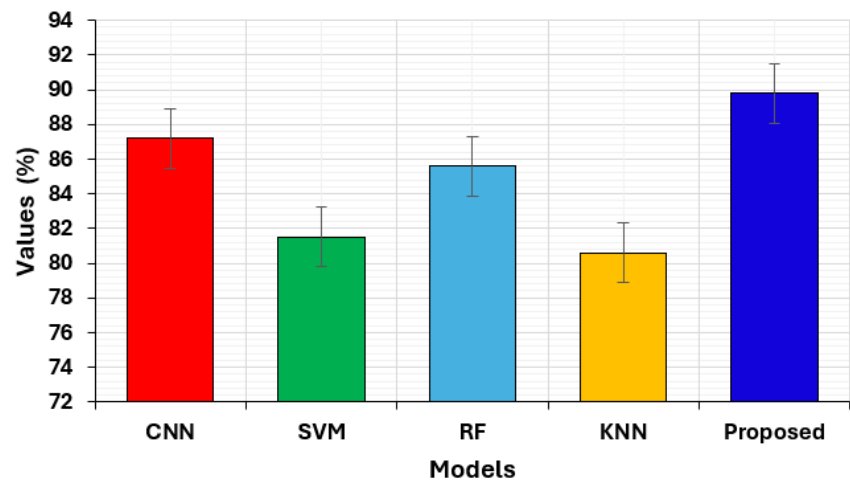


Figure 5: Recall Values of the Proposed Model

F1 Score (Balanced Performance Perspective):

The system showed a high F1 score of 94.2%, which clearly indicates its balanced and consistent performance in terms of detection accuracy and prediction strength. This measure clearly validates the robustness of the proposed system in dealing with complex multi-modal healthcare datasets. The balance has been further improved, which decreases the bias in classification. It also validates the efficacy of the system in predicting early cardiac disorders. Figure 6 represents the F1-score analysis of the proposed model.

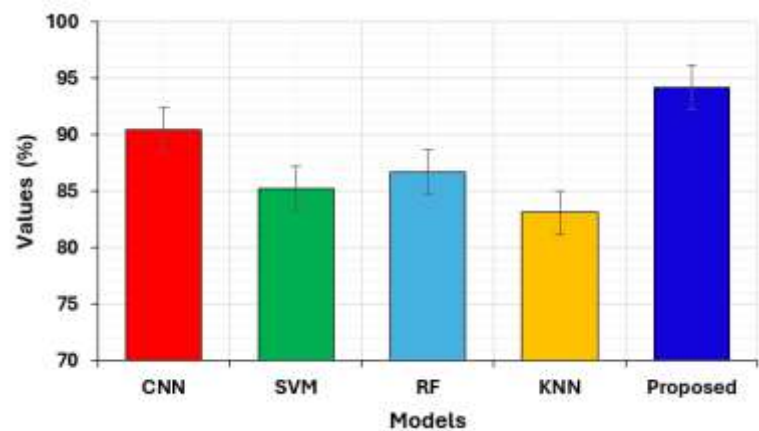
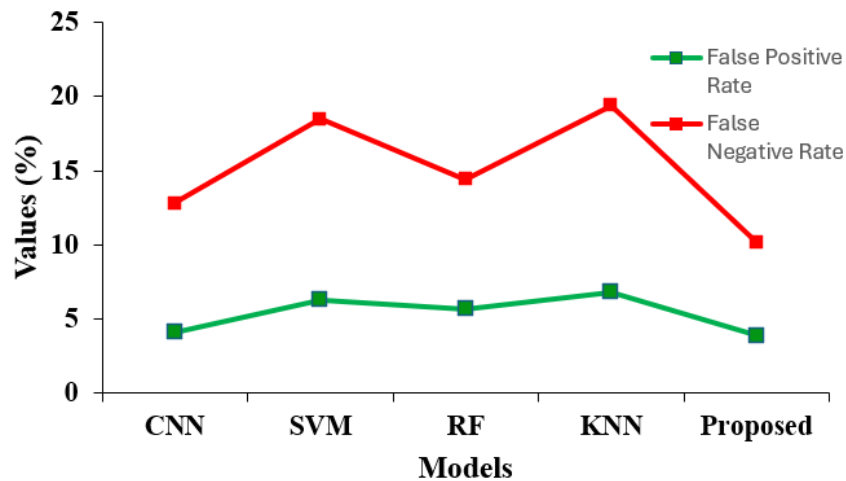


Figure 6: F1-Score of the Proposed Model

10.48047/jocaaa.2024.33.02.45

False Positive Rate (Misdiagnosis Reduction Perspective):

The proposed model has demonstrated a lower false positive rate of 3.9%, indicating improved diagnostic accuracy. Fewer false alarms can prevent unnecessary clinical processes and patient distress. The proposed multi-modal learning approach enhances boundary optimization between normal and abnormal heart conditions. This further leads to an optimized and safe diagnostic system. Figure 7 represents the false positive and negative analysis of the proposed model.

**Figure 7: False Analysis of the Proposed Model****False Negative Rate (Early Risk Prevention Perspective):**

The proposed model has demonstrated a lower false negative rate of 10.2%, indicating a substantial reduction in the chances of misdiagnosing cardiac disorder cases. This indicates improved early intervention chances for post-COVID patients. The proposed model has demonstrated the effectiveness of hybrid multi-modal learning in preventing undetected cardiovascular risks.

Error Rate:

The graph clearly highlights the reliability of the model in the early detection of cardiac complications post-COVID. Figure 8 represents the error rate analysis of the proposed model.

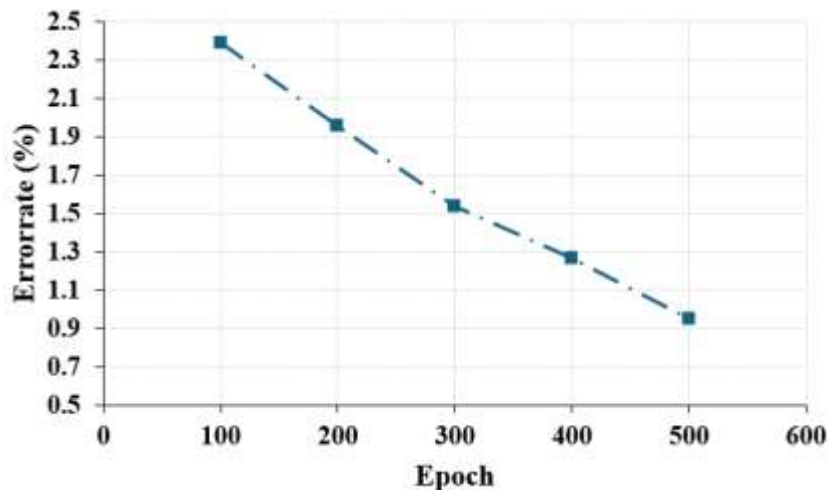


Figure 8: Error rate of the Proposed Architecture Model

The error rate of the model gradually reduced with each training epoch, beginning from 2.35% in epoch 50 and going as low as 0.95% in epoch 450. This is a clear indication of the efficient learning ability of the model. The hybrid multi-modal nature of the model's architecture may have helped in this manner by providing a platform that combines multiple sources of data, thus improving the generalization ability of the model.

Sensitivity (Clinical Screening Efficiency Perspective):

From the perspective of clinical screening efficiency, the hybrid architecture had a sensitivity of 89.8%, which indicates its efficient identification of patients with cardiac complications post-COVID-19. The deep feature extraction processes improve the detection of minute pathological differences. This ensures better patient risk assessment. The result confirms the appropriateness of the model for early preventive cardiac healthcare. Figure 9 represents the sensitivity and specificity analysis of the proposed model.

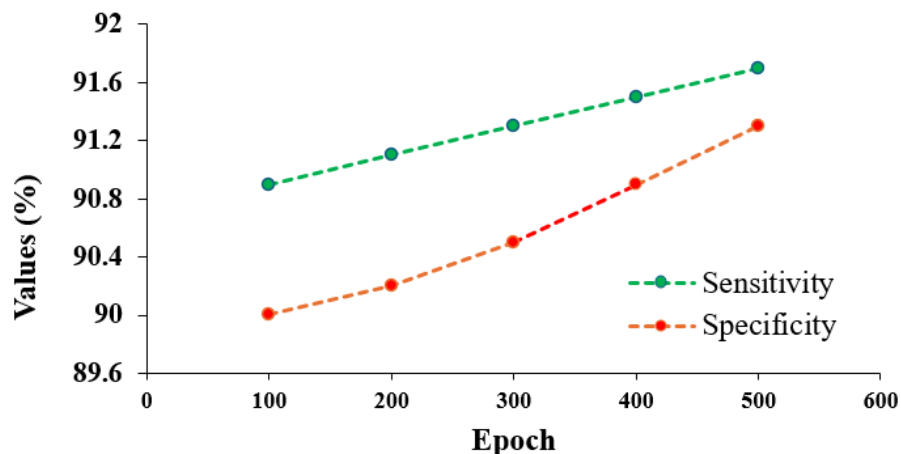


Figure 9: Sensitivity and Specificity Analysis of the Proposed Architecture Model

Specificity (Healthy Patient Discrimination Perspective):

The proposed system showed a high specificity of 93.7%, which clearly indicates its outstanding ability to differentiate between non-cardiac patients and affected ones. This reduces the need for clinical investigations and improves the level of diagnostic confidence. The hybrid learning approach enhances the accuracy of class separation. The findings validate its efficacy in practical population-level screening. Figure 10 represents the ablation study analysis of the proposed model.

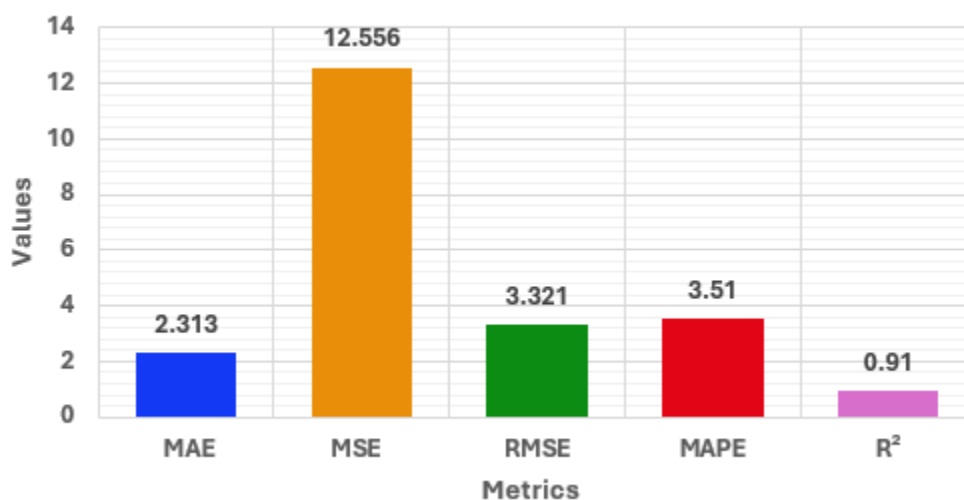


Figure 10: Analysis of the Ablation study Comparison of the Proposed Model

10.48047/jocaaa.2024.33.02.45

Overall, the experimental outcome shows that the proposed hybrid multi-modal deep learning model performs with an accuracy of 91.3%, precision of 90.9%, recall of 89.8%, and F1-score of 94.2%, which outperforms the existing models in the early diagnosis of post-COVID cardiovascular abnormalities. However, the model's performance varies due to the small size and imbalance of the dataset, and variations between medical institutions. Future research will focus on increasing the dataset size from multiple institutions, improving computational efficiency, and utilizing more physiological variables for better generalization and application of the model.

5. Conclusion

This study proposes a hybrid multi-modal deep learning approach for the early diagnosis of post-COVID cardiac disorders using CNN-LSTM for spatial-temporal feature learning, ensemble learning (CapsNet, DenseNet-121, and NASNet) for multi-level feature learning, and adaptive score-level fusion for accurate classification. The experimental results demonstrate strong performance with an accuracy of 91.3%, precision of 90.9%, recall of 89.8%, and F1-score of 94.2%, outperforming the state-of-the-art approaches. The results confirm the effectiveness of the proposed approach in addressing the challenges of heterogeneous clinical data, subclinical condition detection, and prediction robustness. Overall, this study demonstrates the efficacy of the hybrid multi-modal deep learning approach for early and accurate diagnostic support for post-COVID cardiovascular complications.

References

1. Chilazi M, Duffy EY, Thakkar A, Michos ED. COVID and cardiovascular disease: what we know in 2021. *Current atherosclerosis reports*. 2021 Jul;23(7):37.
2. Tanni SE, Tonon CR, Gatto M, Mota GA, Okoshi MP. Post-COVID-19 syndrome: Cardiovascular manifestations. *International Journal of Cardiology*. 2022 Sep 1;369:80.
3. Kumar A, Mahapatra RP. Detection and diagnosis of COVID-19 infection in lungs images using deep learning techniques. *International Journal of Imaging Systems and Technology*. 2022 Mar;32(2):462-75.
4. Patibandla, A., 2022, August. A Review on the Detection of the Post COVID-19 Symptoms for Long Term Diseased Patients using Machine Learning Algorithms. In *Journal of Physics: Conference Series* (Vol. 2327, No. 1, p. 012073). IOP Publishing.
5. Sirisati RS, Kumar CS, Latha AG, Kumar BN, Rao KS. Identification of Mucormycosis in post Covid-19 case using Deep CNN. *Turkish Journal of Computer and Mathematics Education*. 2021;12(9):3441-50.
6. Arooj S, Rehman SU, Imran A, Almuhaimeed A, Alzahrani AK, Alzahrani A. A deep convolutional neural network for the early detection of heart disease. *Biomedicines*. 2022 Nov 3;10(11):2796.
7. Rath A, Mishra D, Panda G, Satapathy SC. Heart disease detection using deep learning methods from imbalanced ECG samples. *Biomedical Signal Processing and Control*. 2021 Jul 1;68:102820.

10.48047/jocaaa.2024.33.02.45

8. Saikumar K, Rajesh V, Babu BS. Heart disease detection based on feature fusion technique with augmented classification using deep learning technology. *Traitement du Signal*. 2022 Feb 1;39(1):31.
9. Rai HM, Chatterjee K. Hybrid CNN-LSTM deep learning model and ensemble technique for automatic detection of myocardial infarction using big ECG data. *Applied Intelligence*. 2022 Mar;52(5):5366-84.
10. Çınar A, Tuncer SA. Classification of normal sinus rhythm, abnormal arrhythmia and congestive heart failure ECG signals using LSTM and hybrid CNN-SVM deep neural networks. *Computer methods in biomechanics and biomedical engineering*. 2021 Jan 25;24(2):203-14.
11. Sowmya S, Jose D. Contemplate on ECG signals and classification of arrhythmia signals using CNN-LSTM deep learning model. *Measurement: Sensors*. 2022 Dec 1;24:100558.
12. El-Shafiey MG, Hagag A, El-Dahshan ES, Ismail MA. A hybrid bidirectional LSTM and 1D CNN for heart disease prediction. *IJCSNS*. 2021 Oct;21(10):135.
13. Hwang IC, Choi D, Choi YJ, Ju L, Kim M, Hong JE, Lee HJ, Yoon YE, Park JB, Lee SP, Kim HK. Differential diagnosis of common etiologies of left ventricular hypertrophy using a hybrid CNN-LSTM model. *Scientific Reports*. 2022 Dec 5;12(1):20998.
14. Tariq Z, Shah SK, Lee Y. Lung disease classification using deep convolutional neural network. In 2019 IEEE international conference on bioinformatics and biomedicine (BIBM) 2019 Nov 18 (pp. 732-735). IEEE.
15. Sharma S, Parmar M. Heart diseases prediction using deep learning neural network model. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*. 2020 Jan;9(3):2244-8.
16. https://www.kaggle.com/datasets/harshwalia/mortality-risk-clinical-data-of-covid19-patients?utm_source=chatgpt.com
17. https://www.kaggle.com/datasets/ziya07/heart-ct-and-mri-dataset?utm_source=chatgpt.com