Volume (1) No (11), 2025
Journal of Global Research in Multidisciplinary Studies (JGRMS)
Review Paper/Research Paper
Available online at https://saanvipublications.com/journals/index.php/jgrms/index

Smart Health Monitoring System with Deep Learning Models for Automated Heart Disease Detection

Mrs. Neha Upadhyay
Assistant Professor
Department of Computer Applications, IIS University
Bhopal (M.P.)
neha.upadhyay887@gmail.com

***Abstract*—Heart disease has remained a major cause of mortality globally and therefore there is the important necessity of early diagnosis and follow-up health studies. Conventional diagnostic methods are also time-consuming, costly and require the interpretation of the expert, and thus they are not suitable in mass screening. In this study, a Smart Health Monitoring System is presented to overcome such problems. This is by incorporating data collection which is enabled by Internet of Things (IoT) and a CNN model to conduct automated detection of cardiac illness. Before the model is trained, the Heart Disease Dataset that is composed of thirteen clinical attributes was preprocessed by data cleansing, normalization, feature extraction, and balancing. As an insistence of its resilience and usefulness in diagnosing heart disease, the proposed CNN model showed good prediction accuracy (ACC) of 99.9, 99.5, 98.4, and F1-score (F1) of 99.4. The CNN model was much better than the more traditional models such as Naive Bayes (NB), Random Forest (RF) and Logistic Regression (LR) in all measures of assessment. All things considered, findings demonstrate that the suggested Smart Health Monitoring System based on CNN has great promise as an intelligent, dependable, and real-time answer to the problems of early cardiac risk assessment and preventative healthcare.**

***Keywords*—Smart Health, Monitoring System, Health Care, Convolutional Neural Network, Heart Disease, Feature Extraction.**

I. Introduction

According to the World Health Organization, noncommunicable diseases (NCDs) cause 41 million deaths each year, or more than 71% of all fatalities. Noncommunicable illnesses claim the lives of 15 million people every year, across all age groups (30–63). 85% of these needless deaths occur in countries with low or medium incomes. Diseases of the cardiovascular system (CVDs) account for the vast majority of NCD cases, with cancer, respiratory illnesses, and diabetes following in decreasing order. These statistics highlight the fact that the global burden of CVDs requires immediate preventive and treatment measures to minimize it [1][2][3]. Diagnostic systems should be the subject of priority in health organizations across the world as the prevalence of heart disease is still increasing worldwide as a major health-related issue. The death toll from cardiovascular illnesses can be reduced with early detection and treatment of IHD, which is still a primary cause of death globally [4].

The heart, a powerful and muscular organ roughly the size of a fist situated about to the left of the midway of the chest, serves as the hub of the circulatory system [5]. Heart illness, often known as cardiovascular disease (CVD), is any ailment that perturbs its regular functioning [6][7][8]. The word "cardiovascular disease" refers to a multitude of medical disorders, including aortic disease, congenital heart disease, pulmonary embolism, deep vein thrombosis, heart failure, cardiomyopathy, strokes, and coronary artery disease.

There are several diagnostic tools that can be used to identify cardiovascular disease. Echocardiograms, stress tests, cardiac MRIs, CT scans, coronary angiography, blood tests, and electrocardiograms are some of these. Nevertheless, these traditional methods have a significant disadvantage, such as possible health hazards, radiation, and specialized knowledge, expensive and labor-intensive processes [9][10]. Preventing cardiovascular disease is significantly easier when risk factors are identified early [11][12]. Deep learning has over the past few years completely transformed the way cardiovascular diseases are predicted due to the significant benefits it has over conventional machine learning techniques [13][14][15]. Healthcare is no exception, as it has experienced a significant change under the influence of IoT, cloud computing, AI, and ML [16][17][18]. Based on the fact that deep learning can reveal the finer details in the complex information and machine learning is more interpretable, the cardiovascular disease diagnosis in healthcare professionals can be more accurate, efficient, and transparent. A possible application of DL models is the identification of latent patterns on large volumes of data, which may result in the identification of heart defects [19][20].

The overview of advanced health monitoring systems has turned the tide in reaction to the increasing rates of cardiovascular conditions and shortcomings of the traditional diagnoses. Wearable sensors with the IoT can also be used in these systems so that they can log physiological measurements constantly, such as the heart rate, blood pressure, and ECG signals [21][22]. Together with the deep learning algorithms, the systems can be used to analyze in real-time and detect potential cardiac anomalies before. Such a combination of IoT and AI does not only ensure that the process of detection is automatic, but much more accurate, less clinical workload, and remote patient monitoring is possible [23]. The results of this study offer an interducture between traditional medical approaches to diagnosing diseases and AI-driven health management systems by presenting a Smart Health Monitoring System that could detect cardiac disease automatically with the help of DL models.

A. *Motivation and Contribution*

The increasing global morbidity of cardiovascular diseases, complemented by the natural limitations of the older and more conventional diagnostic tools, are putting a strain on the contemporary world to find more intelligent, automated and accessible means to healthcare. Traditional methods of heart diseases detection typically require special equipment, special testing, and medical centers and cannot be easily diagnosed in early adulthood particularly in remote or poorly endowed regions. The implementation of smart health monitoring systems that can automatically identify cardiac problems and monitor patients in real-time is becoming a more viable option as deep learning technology and Internet of Things revolutionize at an extremely high rate. The possibility to decrease the mortality rates and enhance the quality of preventative health care provided by the assistance of accurate, economic, and proactive heart health examinations offered by the combination of IoT-based data gathering with deep learning algorithms is what prompts the present study. This study has some important contributions as described below:

- Designing a Smart Health Monitoring System that combines IoT-powered data collection with DL to detect heart diseases automatically.
- Developed and deployed a CNN structure with the capacity to efficiently extract intricate patterns in cardiac data in an attempt to buy out the best predictive results.
- Comparing the proposed CNN model with more traditional ML algorithms (NB, RF, LR) and it was conclusively better than those in regard to ACC.
- The model's robustness, generalizability, and stability were guaranteed by a few assessment criteria and training visualization curves.

B. *Novelty and Justification of the Study*

The originality of the current study is that it uses the DL in order to automate the detection of heart issues in real-time within a Smart Health Monitoring System. The suggested solution depends on IoT-powered data collecting and analysis based on CNN to guarantee continuous monitoring and the early detection of a cardiovascular problem, in contrast to the conventional diagnostic methods that can be laborious, time-consuming, and equipment-dependent. This integration of linked health infrastructure and smart computation is the result of an active, scalable and smart solution to the issues of early diagnosis, clinical load reduction and access to healthcare.

C. *Organization of the Paper*

The following is how the paper is organized. Review previous studies on the diagnosis of heart disease in Section II. Describe the dataset, pre-processing procedures, and model implementation in Section III. In Section IV, report the experimental results and compare them to other studies. Finally, in Section V, summarize the main findings and suggest areas for future research.

II. Literature Review

The development of this study was guided and strengthened by a comprehensive evaluation and analysis of significant research studies on heart disease detection, as shown in Table I.

Venugopal, Natrajan and Sheshadri (2025) focuses on assisting the clinicians in the detection of CHD by advancing techniques for recognition of fetal heart defects using deep learning. The paper presents a computer aided diagnostic tool on the measurement of the Cardio Thoracic Ratio (CTR) which is the predictor of the onset of the CHD. The first step is to train a DL model to automatically detect the heart and thoracic areas and calculate the CT ratio. As it can be seen, the developed deep learning model can help identify the region of interest automatically with an average accuracy of 99.2% [24].

Diwan et al. (2025) DL models like EfficientNet-1D and ResNet-1D are used, along with ML models like XGBoost, SVM, LR, and KNN. F1, ACC, REC, and PRE were among the performance criteria used to evaluate the models. Voting Ensemble Classifier did better than getting the highest overall accuracy of 93.2%. ResNet-1D and EfficientNet-1D are two models of deep learning with an accuracy of 86.6 and 92.2, respectively, and high results in ECG classification. ResNet-1D excelled in detecting normal and arrhythmic signals, while EfficientNet-1D demonstrated superior effectiveness in identifying myocardial infarctions [25].

Alobaid, Bonny and Al–Shabi (2024) applies CNN model to automate the detection of heartbeat diseases (including Artifact, Murmur, Extrastole, and Extrahls) using 1,441 heartbeat sound records as inputs. The model demonstrated accurate heartbeat detection in performance testing, with 98% training accuracy and 92% validation accuracy, correspondingly, for a variety of cardiac conditions. The model demonstrated considerable promise with a peak precision rate of 100% for identifying Extrastole disease. These results are excellent considering the complexity associated with detecting heartbeat diseases [26].

Salau et al. (2024) dedicated to the development and execution of a smart automated system that addresses the obstacles in the detection of cardiac illness. A more precise medical decision support system is necessary for reliable cardiac disease detection, yet there is a lot of room for growth in the current literature. By suggesting a more effective model for cardiac disease diagnosis, this study thereby builds upon previous studies. They have introduced a cutting-edge model with a predicted accuracy of 98.60% for identifying heart illness that makes use of support vector machines and sequential feature selection [27].

Netto and Abraham (2021) proposes DL-based architectures to identify abnormalities in heart sounds. The study establishes five categories: normal (N), four cardiovascular diseases, and unsegmented phonocardiograms. Feature extraction is done using MFCCs, and learning and classification are done using DL techniques like CNN, LSTM, or a combination of the two.. The 392 PCG segments utilized to construct the models are distributed among the several classes. Their CNN, LSTM, and 1DCNN-LSTM accuracy rates were 99.1%, 98.2%, and 99.4%, respectively [28].

Kumar and Reddy (2021) developed a number of programs to automatically detect heart failure. Heart failure diagnostic accuracy is improved during model training and testing using the newly proposed methods. For the purpose of detecting heart failure, this suggested method employs supervised learning, namely the gradient boosting technique. Gradient boosting (GB) is used to train and test the model in the proposed diagnostic system. utilize a gradient boosting classifier to extract the characteristics of cardiac diagnosis. This study goal out to identify cases of heart failure by analyzing the Cleveland Dataset. On par with competing approaches, the suggested methodology achieves a precision of 97.10% [29].

Although the research on the use of DL and ML models to detect heart diseases has gone a long way, there remain various gaps in the research. The majority of the available research is aimed at high classification accuracy but is not applicable to a wide range of datasets and clinical situations in the real world. The models tend to make use of certain modalities, including ECG, PCG, or imaging data alone, as opposed to multimodal data that would be more likely to increase diagnostic accuracy. Explainable AI, which is a critical success factor in clinical adoption, is also not given the required priority as far as model decision interpretation is concerned. Overfitting and decreased resilience when applied in practice are possible outcomes of the tiny or imbalanced datasets used in many experiments. Finally, while high accuracy is reported in controlled experiments, few studies address computational efficiency, real-time applicability, or validation in clinical environments, indicating the need for more comprehensive, interpretable, and clinically validated heart disease detection frameworks.

TABLE I. RECENT STUDIES ON HEART DISEASE DETECTION USING DEEP LEARNING

Author	Proposed Work	Results	Key Findings	Limitations & Future Work
Venugopal, Natrajan & Sheshadri (2025)	Developed a CAD tool for Cardio Thoracic Ratio (CTR) measurement using Mask R-CNN for CHD detection	Automated identification of cardiac/thoracic regions with 99.2% accuracy	Reduces inter- and intra-observer errors in CTR measurement; assists clinicians in CHD detection	Future work could focus on larger datasets and real-time deployment
Diwan et al. (2025)	Employed deep learning models (EfficientNet-1D, ResNet-1D) and ML models (XGBoost, SVM, LR, kNN) for ECG classification	Voting Ensemble achieved 93.2% accuracy; ResNet-1D: 86.6%, EfficientNet-1D: 92.2%	Ensemble models improve ECG classification; ResNet-1D better for normal/arrhythmic signals, EfficientNet-1D better for MI	Could explore hybrid models and larger multi-lead ECG datasets
Alobaid, Bonny & Al–Shabi (2024)	Applied CNN on 1,441 heartbeat sound records for automated heartbeat disease detection	Training accuracy: 98%, Validation accuracy: 92%, Peak precision for Extrastole: 100%	CNN with proper preprocessing allows precise classification of heartbeat diseases	Dataset diversity and the incorporation of real-world noise are potential areas for further research.
Salau et al. (2024)	Using SVMs and sequential feature selection, built a smart automated approach to detect cardiac illness.	Predictive accuracy: 98.60%	SVM with feature selection improves precision in heart disease detection	Scope to enhance system performance further with deep learning integration
Netto & Abraham (2021)	Anomaly detection from heart sounds utilizing MFCC characteristics was accomplished using CNN, LSTM, and 1DCNN-LSTM.	Precision: CNN-99.1%, LSTM-98.2%, 1DCNN-LSTM-99.4%	Deep learning models can classify multiple heart conditions from PCG signals	Future work could include larger datasets and noise-robust methods
Kumar & Reddy (2021)	Developed automated heart failure diagnosis using Gradient Boosting (GB) on Cleveland Dataset	Accuracy: 97.10%	GB improves feature extraction and classification for heart failure detection	Could explore hybrid models or deep learning to further enhance accuracy

III. RESEARCH METHODOLOGY

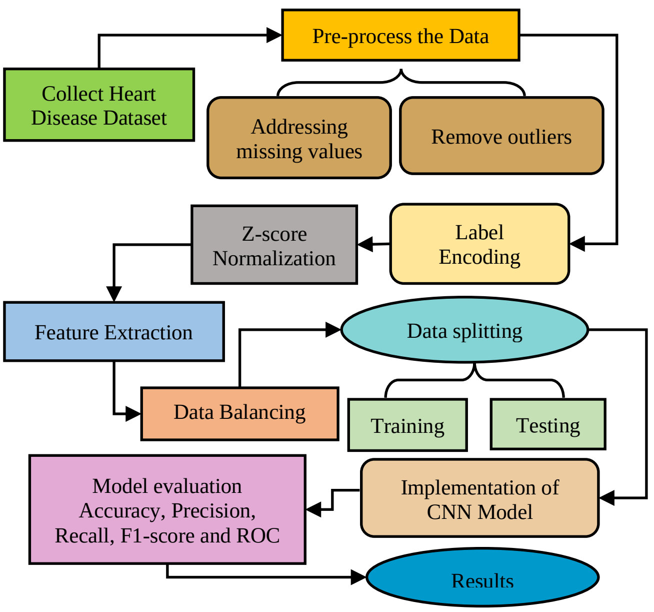

Fig. 1. Proposed Flowchart for Heart Disease Detection using deep Learning

A structured ML pipeline is followed by the suggested methodology for heart disease prediction to achieve trustworthy outcomes. Collecting data and doing any necessary preparation, such as dealing with missing values, eliminating outliers, applying label encoding, and normalizing Z-scores, is the first step. The dataset is divided into training and testing subsets once significant characteristics are found and data balancing techniques are applied. After that, the processed data is used to build a CNN model, which can successfully capture complicated patterns related to heart disease. The Figure 1 displays the flowchart of the proposed method for heart disease prediction are given below:

The following section presents a detailed description of each step in the proposed methodology:

A. *Data Gathering and Analysis*

The Heart Disease Dataset is used in this research. Using information from one thousand data samples over thirteen attributes—each of which represents a possible risk factor—it is useful for tasks related to the prediction and categorization of cardiovascular illnesses. The following data visualizations, which included heatmaps and bar graphs, were utilized to look at feature correlations, attack dispersion, etc.

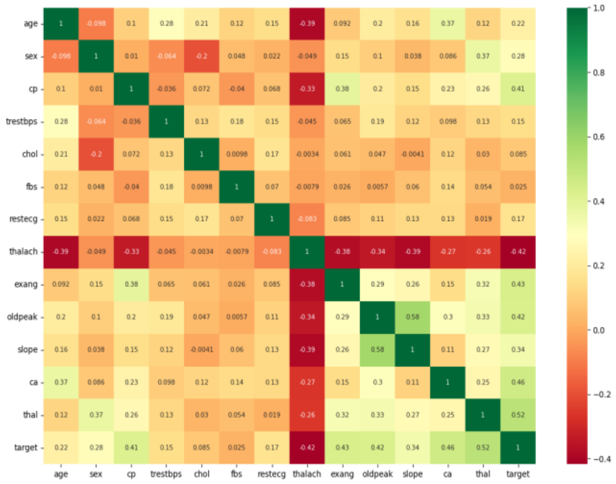

Fig. 2. Heatmap Distribution of the Dataset Features

Figure 2 The heatmap illustrates the correlation matrix of the features within the heart disease dataset, highlighting relationships between clinical variables and the target outcome. Strong positive correlations are shown in darker green, while negative or weaker correlations appear in red to yellow shades. The aim is strongly connected with thalach (maximum heart rate obtained), while it is adversely correlated with exang (exercise-induced angina) and old peak (ST depression). There is less of a correlation between age, cholesterol, and sBP when patient is at rest. Overall, the heatmap can help in selection and interpretation of a predictive model by revealing the relationship and association between different variables in the patients and the likelihood of cardiovascular disease.

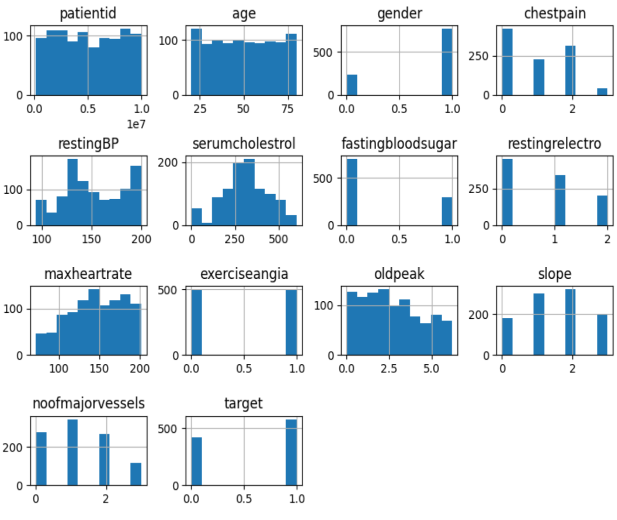

Fig. 3. Histogram Distribution of the Dataset Features

Figure 3 shows the heart disease dataset's histogram plots, which display the features' distribution graphically and highlight the dispersion and variability of each attribute. Gender is binary, with a larger count in one group, while age seems to be pretty evenly distributed between 30 and 75 years. There is a categorical distribution for the categories of chest discomfort, resting electrocardiographic data, slope, and number of main vessels. The distribution of serum cholesterol is right-skewed, although the ranges of resting blood pressure, maximal heart rate, old peak, and serum cholesterol are completely varied. There are aspects that generally have binary values, such as exercise-induced angina and fasting blood sugar. This implies that the majority of values may be assigned to a mere one. The target variable is handy to classify, as it determines the classified group with cardiac disease and the other group without cardiac disease. The plots provide critical input towards predictive modelling as they depict a combination of continuous and categorical data with different distributions.

B. Data Pre-processing

Handling missing values and outliers the Heart Disease Dataset preprocessing procedure included several key processes, including data normalization, label encoding of categorical variables, and detection and treatment.. Based on these steps, the dataset was cleaned and structured in a way allowing the training of DL models on it. Proper preprocessing had the benefits of improving the overall efficiency of the process of predicting heart diseases, reducing bias, and improving the accuracy of the models:

- **Addressing missing values:** Time series data recovery and repair algorithms need to consider time dependencies and preserve the dataset structure. Common methods include removing data, filling in gaps in data, estimating missing values, and imputation based on models.
- **Detecting and managing outliers:** Outlier treatment includes removing the points, imputing them with central values (like the median), capping them at a defined upper/lower limit, or applying transformations to reduce their influence.
- **Label Encoding:** Machine learning makes use of label encoding, a data preparation method, to transform numerical data from categorical data. This change is crucial because many ML algorithms can only process numerical data and not categorical data in its raw form.

C. Z-score Normalization

Data normalization seeks to alter or standardize data in order to make it more consistent. Data normalization methods include scaling, min-max normalization, and z-score normalization. Many others are used often. Included among the standardization methods employed in this inquiry was the z-score normalization procedure, which has a mean of 0 and a standard deviation of 1. Values near the mean with a standard deviation of one are altered by this scaling approach. Additionally, Equation (1) defines z-score normalization.

$$E' = \frac{E-\bar{M}}{\sigma_M} \quad (1)$$

$\bar{M}$ is the mean, while σ_M is the standard deviation. For each data entry, E' and E are the new and old values, respectively.

D. Feature Extraction

Data scientists and ML experts use feature extraction as a pre-processing step to simplify and improve the quality of complex raw data. The aim is to improve model performance, interpretation, and efficiency while reducing computational expenses and data dimensionality by creating new, reduced features that capture the essential information in the original data.

E. Data Balancing using Oversampling

The fundamental principle of data balancing, which is the solution to the skewed distribution of classes in data sets, has remained unchanged, the definition has shifted to include more advanced, performance-oriented, and resource-conscious techniques. Oversampling Data balancing by oversampling promotes the representation of a minority population within an imbalanced dataset by sampling the existing data points (or creating new ones).

F. Data Splitting

The dataset was partitioned into training and assessment subsets to measure its efficiency. In particular, when building the model and estimating its parameters, used 80% of the data, while testing and evaluating the model received the remaining 20%.

G. Proposed Convolutional Neural Network (CNN) Model

The purpose of this paper is to propose a CNN, a DL approach, for the identification of heart disease. The pattern recognition and image processing industries frequently employ CNNs, DL techniques. A CNN is built using convolutional, pooling, and fully linked layers. The convolutional layer simplifies the input parameters while the pooling layer converts the images to numbers. An excellent deep learning method for picture processing and recognition is the CNN. Some of its components include convolutional, pooling, and fully linked layers. CNNs attempt to reproduce the visual processing capabilities of the human brain. Hierarchical patterns and spatial linkages in photographs are areas where they truly shine.t enables us to identify the size of the output of a convolutional layer. The length of output is 5, in this case, Overall, the output length is as follows Equation (2),

$$\text{Output size} = \text{nx} = 2\text{P} - \text{nhS} + 1, \tag{2}$$

The magnitude of the output can be expressed as nx + 2 P − n h S + 1, where nx is the input signal length and nh is the filter length.

The mathematical operation known as convolution (Conv_Op) has numerous applications in computer vision, signal processing, and image processing. It is a mathematical operation that is used to add two signals together or two functions that are used to generate a third signal representing the effect of one signal upon the other, but weighted by the shape of the other signal. Convolution is applied to images in computer vision and it is a method of image feature extraction with CNNs.

The mathematical Equation (3) definition of the convolution operation is as follows:

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m] \tag{3}$$

In this case, f and g are discrete or continuous functions, and n is the index of position or time of the output signal. The $*$ is used to represent the convolution operation. The above Equation (4) can be expressed as when the input signals are discrete, the equation becomes as follows.

$$(f * g)[n] = \sum_{m=-\infty}^{\infty} f[m]g[n-m]\Delta m \tag{4}$$

In this case, f and g are either discrete or continuous functions and n is the position or time index of the output signal. The convolution image is represented by the character $*$. The equation above can then be expressed as when the input signals are discrete.

$$(f * m)(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau \tag{5}$$

where t is the index of time of the output signal show in Equation (5).

H. Evaluation Metrics

The proposed model was measured with the help of various metrics. A confusion matrix was used to illustrate the categorization results, stating the correct and incorrect predictions made in each of the classes. This matrix was used to obtain all the four True Negative values- TP, FP, TN and FN. The following table shows the results of applying these variables to compute important evaluation metrics: REC, ACC, PRE, and F1:

Accuracy: The degree to which the outputs were accurately predicted by the trained model in comparison to the entire dataset (input samples). The Equation (6) is -

$$Accuracy = \frac{\text{TP}+TN}{\text{TP+Fp+TN+FN}} \tag{6}$$

Precision: A model's prediction accuracy is determined by dividing the number of correctly predicted positive cases by the total number of positive instances. Precision is assumed. Equation (7) expresses the classifier's accuracy in predicting positive classifications-

$$Precision = \frac{\text{TP}}{\text{TP+FP}} \tag{7}$$

Recall: There should have been a positive ratio between all cases that were projected to be positive and the actual number of positive predictions. Its mathematical representation is Equation (8)-

$$Recall = \frac{\text{TP}}{TP+FN} \tag{8}$$

F1 score: This combination, which resembles a harmonic mean of the two, balances out both REC and PRE. It ranges from 0 to 1. In terms of math, it is expressed as Equation (9)-

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{9}$$

Receiver Operating Characteristic Curve (ROC): The ROC plots, for a set of decision cut-off points, the ratio of successfully categorized cases to those that were wrongly classified. FPR is equivalent to 1-specificity, although TPR is frequently referred to as recall or sensitivity.

IV. RESULTS AND DISCUSSION

The experimental setup was implemented on a Dell LATITUDE E6430 system (Core i7, 4GB RAM) using Google Colab as the development environment. The Heart Disease Dataset was used to train and assess the proposed model. Several highly important critical performance metrics, including ACC, PRE, REC, and F1, were used to assess its performance. Table II summarizes the data and shows how well the suggested model performs in categorization. With a 99.9% ACC rate, 99.5% PRE rate, 98.4% REC rate, and 99.4% F1, CNN was determined to be very efficient and dependable in identifying instances of cardiac disorders.

TABLE II. CLASSIFICATION RESULTS OF THE PROPOSED MODEL HEART DISEASE DETECTION USING HEART DISEASE DATASET

Matrix	Convolutional Neural Network (CNN)
Accuracy	99.9
Precision	99.5
Recall	98.4
F1-score	99.4

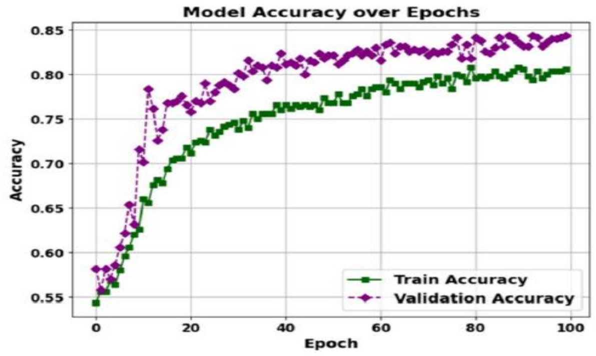

Fig. 4. Accuracy curve for the CNN Model

In both datasets, Figure 4 shows the model's ACC trend throughout 100 epochs of training and validation data. The ACC in training (green squares), and validation accuracy (purple diamonds) are steeply increasing at the first few epochs and improve between approximately 0.55 to more than 0.85 in the first 20 epochs. Both accuracies increase at a slow pace as training progresses and start to level off between epoch 80 and 100. The model is considered well-generalized because learning is accomplished effectively when the average validation accuracy is higher than the training accuracy throughout the procedure, which reaches a maximum value of 0.84.

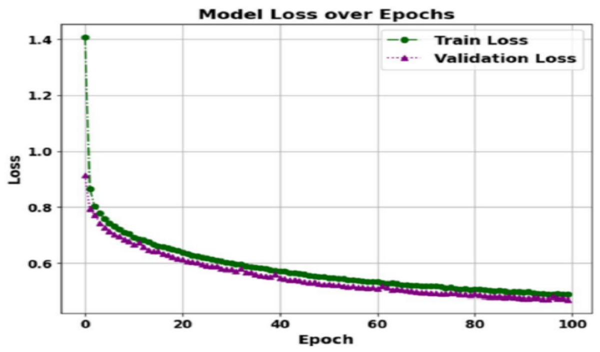

Fig. 5. Loss curve for the CNN Model

The model's training and validation loss, expressed in terms of 100 epochs, is shown in Figure 5. Over the first twenty epochs, there is a precipitous decline in both the training loss (green circles) and validation loss (purple triangles), decreasing from 1.4 to less than 0.8. The training losses keep reducing and stabilize between the 80th and 100th epochs as the training continues, and the two curves come closer to 0.5. The steady decrease and near correspondence of training and validation losses demonstrate that the training was successful, there was little overfitting, and the models remained stable during the training.

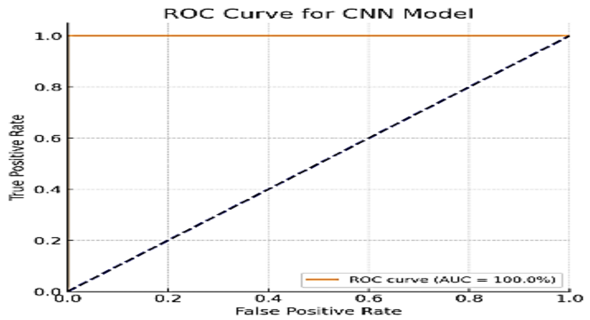

Fig. 6. ROC Curve for CNN Model

Figure 6 shows the CNN model's ROC curve, which shows how well it classified data. The ROC curve (orange) is at the top of the plot and this means that the AUC is 100%. It shows the extraordinarily high positive and negative rate distinction of the model with a TP of 1 and a FP of 0. The basis of random classification is the diagonal dashed line, and the fact that the curve of the model is far above it prove the excellent predictive performance and discriminative competence.

A. *Comparative Analysis*

Table III details the results of comparing the accuracy of the proposed CNN model to those of other current models in order to ascertain its efficacy. A conventional ML model, NB achieved 83% ACC, 81% PRE, 85% REC, and 85% F1. With an F1 of 94%, an ACC of 90.4%, a PRE of 93.1%, and a REC of 95.2%, RF proved to be the superior method. LR performed even better, with a 96.1% F1, 95.7% REC, 96.5% PRE, and 95.5% ACC. With a 99.9% ACC rate, 99.5% PRE rate, 98.4% REC rate, and 99.4% F1, the CNN model outperformed all other models in the competition for recognizing heart disease in the dataset.

TABLE III. COMPARISON OF DIFFERENT MACHINE LEARNING MODELS FOR HEART DISEASE DETECTION ON HEART DISEASE DATASET

Model	Accuracy	Precision	Recall	F1-score
NB[30]	83	81	85	85
RF[31]	90.4	93.1	95.2	94
LR[32]	95.5	96.5	95.7	96.1
CNN	99.9	99.5	98.4	99.4

The CNN model was applied in an automated heart disease detection Smart Health Monitoring System using deep learning as an efficient method of analyzing complex cardiac data. It had a great level of accuracy of 99.9% which indicated that it is very reliable and accurate in detecting cases of heart diseases. CNN model is strong and efficient when compared to the current models and this means that the model performs better. These results indicate that it is most likely to be an important tool in real-time health tracking and appropriate diagnosis.

V. CONCLUSION AND FUTURE STUDY

Cardiovascular diseases remain one of the major health issues in the global arena, and promptness in diagnosis plays a central role in reducing the mortality rates and improving patient survival. This study proposed a Smart Health Monitoring System, which constitutes the combination of IoT data-gathering and CNN that are designed to automatically identify cardiac illness. The methodology provided in this paper showed good predictive accuracy, which was able to effectively model complex trends in cardiac data, which were better than the standard ML implementation of NB, RF, and LR. The system offers a believable system of real-time tracking which offers the chance to diagnose at an early stage and give constant health examination and active intervention. The results suggest the potential of the application of deep learning and intelligent health technologies to provide quality, scalable, and accessible cardiac care. The system can be enhanced in future use by increasing the size and variety of datasets, multi-modal data streams such as ECG signals, wearable sensor data and taking into account hybrid deep learning architectures that can achieve improved generalization. Integration with cloud computing and edge AI can further enable low-latency, real-time processing for remote and resource-constrained environments. Also, explainable AI methods may enhance interpretability and trust to implement it clinically, and longitudinal monitoring may facilitate predictive analytics to preventive care and individualized treatment plans.

REFERENCES

[1] A. Damasceno, "Noncommunicable disease," in *The Heart of Africa*, Wiley, 2016, pp. 155–157. doi: 10.1002/9781119097136.part5.

[2] R. P. Mahajan, "Development of Predictive Models for Early Detection of Alzheimer's Disease Using Machine Learning," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 2, pp. 115–123, 2025.

[3] S. Pandya, "A Machine Learning Framework for Enhanced Depression Detection in Mental Health Care Setting," *Int. J. Sci. Res. Sci. Eng. Technol.*, vol. 10, no. 5, pp. 356–368, Oct. 2023, doi: 10.32628/IJSRSET2358715.

[4] H. Gajjar, U. Kora, and V. Gondaliya, "Telemedicine with Automated ECG Analysis," *Int. J. Eng. Res. Technol.*, vol. 2, no. 10, pp. 2253–2255, 2013.

[5] S. Thangavel, "Method for Real-Time Image Processing in IoT-Based Smart Home Security Systems," 2024

[6] A. Baccouche, B. Garcia-Zapirain, C. C. Olea, and A. Elmaghraby, "Ensemble Deep Learning Models for Heart Disease Classification: A Case Study from Mexico," *Information*, vol. 11, no. 4, p. 207, Apr. 2020, doi: 10.3390/info11040207.

[7] S. P. Kalava, "The Role of AI in Reinventing Hospitality Safety Measures After COVID-19," *J. Artif. Intell. Cloud Comput.*, vol. 3, no. 1, pp. 1–3, Feb. 2024, doi: 10.47363/JAICC/2024(3)E153.

[8] S. Pandya, "A Machine and Deep Learning Framework for Robust Health Insurance Fraud Detection and Prevention," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 3, no. 3, pp. 1332–1342, Jul. 2023, doi: 10.48175/IJARSCT-14000U.

[9] C. W. Tsao *et al.*, "Heart Disease and Stroke Statistics—2022 Update: A Report From the American Heart Association," *Circulation*, vol. 145, no. 8, Feb. 2022, doi: 10.1161/CIR.0000000000001052.

[10] S. P. Kalava, "Digital Transformation in the COVID Era: Evolution, Impact, and Future Pathways," *Int. J. Sci. Res.*, vol. 12, no. 4, p. 7, 2023.

[11] K. Mridha, S. Ghimire, J. Shin, A. Aran, M. M. Uddin, and M. F. Mridha, "Automated Stroke Prediction Using Machine Learning: An Explainable and Exploratory Study With a Web Application for Early Intervention," *IEEE Access*, vol. 11, 2023, doi: 10.1109/ACCESS.2023.3278273.

[12] A. Balasubramanian, "Intelligent Health Monitoring: Leveraging Machine Learning and Wearables for Chronic Disease Management and Prevention," *Int. J. Innov. Res. Eng. Multidiscip. Phys. Sci.*, vol. 7, no. 6, pp. 1–13, 2019.

[13] Z. K. D. Alkayyali, S. A. Bin Idris, and S. S. Abu-Naser, "A Systematic Literature Review Of Deep And Machine Learning Algorithms In Cardiovascular Diseases Diagnosis," *J. Theor. Appl. Inf. Technol.*, vol. 101, no. 4, 2023.

[14] K. Mohapatra, J. Mishra, S. R. Pattanaik, A. Pati, A. Panigrahi, and B. Sahu, "EFSMLHA: Ensembled Feature Selected Machine Learning Hybrid Approaches for Heart Disease Prediction," in *2025 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, 2025, pp. 1–6. doi: 10.1109/APSIT63993.2025.11086260.

[15] V. K. Singh, D. Pathak, and P. Gupta, "Integrating Artificial Intelligence and Machine Learning into Healthcare ERP Systems: A Framework for Oracle Cloud and Beyond," *ESP J. Eng. Technol. Adv.*, vol. 3, no. 2, pp. 171–178, 2023, doi: 10.56472/25832646/JETA-V3I6P114.

[16] R. G. Nadakinamani *et al.*, "Clinical Data Analysis for Prediction of Cardiovascular Disease Using Machine Learning Techniques," *Comput. Intell. Neurosci.*, 2022, doi: 10.1155/2022/2973324.

[17] S. R. Sagili, C. Goswami, V. C. Bharathi, S. Ananthi, K. Rani, and R. Sathya, "Identification of Diabetic Retinopathy by Transfer Learning Based Retinal Images," in *2024 9th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2024, pp. 1149–1154. doi: 10.1109/ICCES63552.2024.10859381.

[18] S. Thangavel, "Real-Time Soil Health Monitoring Device with ML," 1224297, 2024

[19] K. Ainapure, K. Sumit, and S. M. Pattanshetty, "A study on implementation of national programme for prevention and control of cancer, diabetes, cardiovascular diseases and stroke in Udupi district, Karnataka," *Int. J. Community Med. Public Heal.*, 2018, doi: 10.18203/2394-6040.ijcmph20182163.

[20] R. P. Mahajan and N. Jain, "Deep Learning Techniques for Identification of Parkinson's Disease based on MRI Data: Novel Diagnostic Framework," in *2025 International Conference on Intelligent Computing and Knowledge Extraction (ICICKE)*, Jun. 2025, pp. 1–7. doi: 10.1109/ICICKE65317.2025.11136577.

[21] S. S. S. Neeli, "Heart Disease Prediction For A Cloud-Based Smart Healthcare Monitoring System Using Gans And Ant Colony Optimization," *Int. J. Med. Public Heal.*, vol. 14, no. 4, 2024, doi: : 10.70034/ijmedph.2024.4.223.

[22] S. Pandya, "Integrating Smart IoT and AI-Enhanced Systems for Predictive Diagnostics Disease in Healthcare," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 10, no. 6, Dec. 2024, doi: 10.32628/CSEIT2410612406.

[23] M. A. Mostafiz, "Machine Learning for Early Cancer Detection and Classification : AI- Based Medical Imaging Analysis in Healthcare," *Int. J. Curr. Eng. Technol.*, vol. 15, no. 3, pp. 251–260, 2025, doi: https://doi.org/10.14741/ijcet/v.15.3.7.

[24] S. T. Venugopal, S. Natrajan, and S. Sheshadri, "Cardio Thoracic Ratio for Congenital Heart Disease detection using Deep Learning," in *2025 International Conference on Biomedical Engineering and Sustainable Healthcare (ICBMESH)*, 2025, pp. 1–5. doi: 10.1109/ICBMESH66209.2025.11182212.

[25] S. Diwan, N. Yadu, D. Singh, M. Chilamwar, and M. Sahu, "Digitization and Classification of Electrocardiograms for Automated Cardiovascular Disease Diagnosis Using Machine Learning and Deep Learning Models," in *2025 International Conference on Ambient Intelligence in Health Care (ICAIHC)*, 2025, pp. 1–6. doi: 10.1109/ICAIHC64101.2025.10956735.

[26] A. Alobaid, T. Bonny, and M. Al–Shabi, "Automating Heartbeat Disease Detection Using Advanced Machine Learning Algorithm," in *2024 Advances in Science and Engineering Technology International Conferences (ASET)*, 2024, pp. 1–7. doi: 10.1109/ASET60340.2024.10708735.

[27] A. O. Salau, T. A. Assegie, G. Chhabra, K. Kaushik, and S. L. Braide, "Heart Disease Detection Model Using Support Vector Machine with Feature Selection," in *2024 2nd International Conference on Advancement in Computation & Computer Technologies (InCACCT)*, 2024, pp. 204–207. doi: 10.1109/InCACCT61598.2024.10550988.

[28] A. N. Netto and L. Abraham, "Detection and Classification of Cardiovascular Disease from Phonocardiogram using Deep Learning Models," in *Proceedings of the 2nd International Conference on Electronics and Sustainable Communication Systems, ICESC 2021*, 2021. doi: 10.1109/ICESC51422.2021.9532766.

[29] K. L. Kumar and B. E. Reddy, "Heart Disease Detection System Using Gradient Boosting Technique," in *Proceedings - 2021 International Conference on Computing Sciences, ICCS 2021*, 2021. doi: 10.1109/ICCS54944.2021.00052.

[30] Khan *et al.*, "Heart Disease Detection: A Comprehensive Analysis of Machine Learning, Ensemble Learning, and Deep Learning Algorithms," *Nano Biomed. Eng.*, vol. 16, no. 4, pp. 677–690, 2024, doi: 10.26599/NBE.2024.9290087.

[31] H. J. Suleiman, I. R. A. Hamid, and O. R. Olaniran, "Smart Health Monitoring for Predicting Heart Disease using IoT-Fog-Cloud Computing Model," *Eng. Technol. Appl. Sci. Res.*, vol. 15, no. 3, pp. 22565–22572, 2025, doi: 10.48084/etasr.10048.

[32] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases," *Diagnostics*, vol. 14, no. 2, 2024, doi: 10.3390/diagnostics14020144.

