Enhancing Heart Disease Detection Using Convolutional Neural Networks and Classic Machine Learning Methods

Sri Hasta Mulyani[1]*, Nurhadi Wijaya[2], and Fike Trinidya[3]

[1-3] Department of Informatics, Universitas Respati Yogyakarta, Yogyakarta, Indonesia

* Correspondence: hasta@respati.ac.id

Abstract: This study addresses the problem of heart disease detection, a critical concern in public health. The research aims to compare the performance of Convolutional Neural Networks (CNN) with conventional machine learning algorithms in diagnosing heart disease using a dataset comprising 14 features. The primary objective is to determine whether CNNs can provide more accurate and reliable results than traditional techniques. The research employs rigorous preprocessing, normalizing relevant features, and splits the dataset into an 80-20 training-testing split. The model is trained for 300 epochs with a batch size of 64, and performance evaluation is conducted using confusion matrices and classification reports. The results reveal that the CNN model achieved a remarkable accuracy of 100%, demonstrating its potential to outperform conventional machine learning algorithms. These findings emphasize the significance of deep learning techniques in improving heart disease diagnostics, although further research is needed to optimize CNN models and address interpretability concerns for practical implementation in healthcare settings.

Keywords: Convolutional Neural Networks; Diagnostic Accuracy; Healthcare Applications; Heart Disease Detection; Machine Learning Algorithms.

1. Introduction

Early detection of heart disease can contribute significantly to public health by enabling timely treatment and reducing mortality. Machine learning techniques, such as Logistic Regression (LR), K-Nearest Neighbor, Decision Tree, Naive Bayes, Random Forest, and Support Vector Machine, have been applied to develop heart disease diagnosis models [1]. These models use various patient characteristics, such as age, chest pain, blood pressure, gender, cholesterol, and heart rate, to identify high-risk individuals and distinguish between patients with and without heart disease [2]. In addition, computer vision and machine learning methods have been used to analyze irises for early detection of heart disease [3]. The proposed system achieved high accuracy in detecting potential heart disease using iridology [4]. By accurately detecting heart disease and supporting doctors in decision-making, this machine learning approach can improve patient outcomes and contribute to public health [5].

Early detection of heart disease is an essential concern in the medical field as it can prevent fatal outcomes and save lives [6]. Heart disease is the leading cause of death worldwide, and detecting it early allows for timely intervention and treatment [7]. By analyzing risk factors and using machine learning techniques, medical professionals can improve decision-making and enhance early detection of heart disease. Deep learning networks and remote diagnostic systems can also support doctors in diagnosing heart disease more effectively. Early detection allows medical practitioners to implement appropriate interventions, such as lifestyle changes, medication, or surgery, to prevent further

complications and improve patient outcomes. Therefore, early detection of heart disease is critical to reducing mortality rates and improving the overall health of individuals.

Heart disease detection methods include analyzing electrocardiogram (ECG) signals, listening to heart sounds, and using medical tests and imaging techniques such as CT scans. Machine learning (ML) and deep learning (DL) models have been used for heart disease classification from ECG signals, with the Generative Adversarial Network (GAN) model and GAN-LSTM ensemble model showing promising results in handling imbalanced data and achieving higher accuracy [8]. Deep learning techniques have also been applied to analyze heartbeat audio recordings, combining convolutional and recurrent neural network models to improve accuracy [9]. Feature extraction methods using ECG signals, such as discrete wavelet transform (DWT) and SVM classifiers, are applicable in early heart disease detection [10]. Machine learning algorithms such as Random Forest (RF), SVM, LR, and KNN have been used with feature selection methods to diagnose heart disease, with RF and RFFS yielding the highest accuracy [11]. Data mining methods, including decision trees, KNN classifiers, naive Bayes, random forests, and SVM, have also been used for heart disease diagnosis, with SVM showing the best performance [12].

Convolutional Neural Network (CNN) was chosen as the primary approach in this study due to its potential for early detection of heart disease and its ability to analyze medical images [13]. CNN architecture can assist in interpreting medical images related to heart disease by utilizing deep learning techniques to process large data sets and accurately predict heart disease [14]. CNN models have been successfully used for left ventricle segmentation in cardiac MRI, providing clinical measurements [15]. The study proposed a CNN-based Fully Connected layer architecture for automatic classification of ECG signals into specific classes, achieving high accuracy results [16]. CNNs have also been widely used in medical image analysis, enabling significant advances in computer-aided diagnosis in image classification, segmentation, detection, and more.

CNNs have several advantages over conventional machine learning techniques in the context of heart disease detection. First, CNNs can automatically learn relevant features from raw data, such as ECG signals, without manual feature engineering [17]. This allows CNNs to capture complex patterns and relationships in the data, improving disease classification accuracy. In addition, CNNs are highly effective in handling high-dimensional data, such as ECG signals, due to their ability to learn hierarchical representations [18]. This allows CNNs to effectively extract and utilize spatial and temporal information from the data, which is crucial for accurate heart disease detection. Furthermore, CNNs have demonstrated superior performance to conventional machine learning techniques regarding classification accuracy, achieving high accuracy in training, validation, and testing [19]. Overall, using CNNs in heart disease detection offers improved accuracy, automatic feature learning, and effective handling of high-dimensional data.

Applying conventional machine learning techniques such as LR, KNN, Decision Tree, Naive Bayes, Random Forest, and SVM can help compare performance with the CNN approach. Various studies have used these techniques to analyze and predict different outcomes. For example, in a study by Uma Pavan Kumar et al. the authors discuss the use of these algorithms and compare their performance [20]. Another study by Ali Shehadeh et al. proposed using Modified Decision Tree regression, LightGBM, and XGBoost to predict the salvage value of construction equipment [21]. In addition, Nikolai Stepanov et al. utilized various machine learning techniques, including SVMs, to predict LTE network edge traffic [22]. These studies demonstrate the effectiveness of conventional machine learning techniques in different domains and highlight their potential for performance comparison with CNN approaches to yield effective results [23].

This study aims to evaluate and compare the performance of CNN with five conventional machine learning techniques, namely LR, KNN, Decision Tree, Naive Bayes, Random Forest, and SVM in heart disease detection. Using a dataset of 12 heart disease-related features, this study will measure the performance of each method using metrics such as Confusion Matrix and Classification Report. The main objective of this research is to provide an in-depth understanding of the effectiveness of each approach in detecting heart disease based on its features. The results of this comparison are expected to provide new insights regarding the most suitable approach for heart

disease detection based on the characteristics of the given dataset, as well as contribute to the development of more accurate heart disease detection methods in the future.

2. Literature Review

The LR approach has been applied in heart disease detection to improve accuracy and speed up the diagnosis process. It is a simple algorithm that has shown good classification ability in previous studies [24]. LR is used to extend the concept of Linear Regression with categorical dependent variables. Compared to other algorithms, such as Random Forest, it has shown greater accuracy in heart disease classification [25]. The advantages of LR in this context include its simplicity, good classification ability, and ability to handle categorical data [26]. However, it also has limitations, such as the assumption of linearity between the independent variable and the log-odds of the dependent variable and the potential for overfitting if the model is too complex [27].

SVM is used in heart disease detection by utilizing machine learning algorithms to predict and classify heart conditions. SVM is a supervised machine learning algorithm that can predict heart disease based on features or symptoms. It is applied in various studies to improve the diagnosis and detection of heart diseases. SVM classifiers are trained and tested using data sets such as the Z-Alizadeh Sani Heart Dataset (HD) [28]. The SVM algorithm uses a feature space, including age, chest pain, blood pressure, gender, cholesterol, and heart rate, to classify heart disease data [29]. Features extracted from ECG signals, such as frequency information obtained through the short-term Fourier transform algorithm, are used as input for the SVM classifier [30]. The model's performance can be improved by tuning the parameters of the SVM classifier [31]. Overall, SVM is a valuable tool in heart disease detection and classification, contributing to early diagnosis and timely intervention.

The Decision Tree Classifier is a commonly used approach in heart disease detection. A tree-like model simulates a hierarchical structure to analyze complex data. The classifier is built based on risk factors such as age, gender, smoking habits, physical activity, obesity, diabetes, stress, and diet [32]. This approach addresses the complexity problem by extracting high-level knowledge from raw data using mining techniques [33]. It simplifies the decision-making process by representing the relationship between risk factors and heart disease in a visual and interpretable manner [34]. To address the issue of overfitting, decision tree classifiers can be optimized using techniques such as pre-trimming, which involves resampling the data set to eliminate overfitting [35]. This helps improve the model's generalization ability and prevents it from memorizing the training data too closely.

KNNs are applied in heart disease classification by using machine learning techniques to predict potential patients. The KNN algorithm has demonstrated its effectiveness, achieving an accuracy rate of 82.03% in classification, and obtaining a performance score of 0.935 when measured using AUC [36]. In the context of heart disease, KNN has been used to predict the likelihood of a person having heart disease based on various attributes. For example, one study achieved an accuracy rate of 86.95% using KNN with 12 attributes [37]. The selection of the K parameter in KNN affects the classification results by determining the number of nearest neighbors considered when making predictions. Different K values can lead to different accuracy rates. For example, in the same study, the KNN accuracy rate varied between 89.29% for a sample size of 20 data points and 96.66% for 299 data points [38].

Random Forest Classifier is advantageous as an ensemble method in heart disease detection because it can handle continuous and categorical data, model nonlinear relationships, and adapt hyperparameters for improved performance [39]. Combining multiple decision trees in the Random Forest Classifier improves accuracy by creating a diverse set of decision trees and simultaneously determining the optimal number of trees [40]. This approach generates different training sets with other samples and features to train each tree, improving the performance of random forests and increasing prediction accuracy [41]. The decision tree ensemble in the Random Forest Classifier outperformed individual classifiers and other ensemble methods, such as random forests and rotational forests, achieving high accuracy scores on heart disease datasets [42].

Machine learning methods such as LR, SVM, Decision Tree Classifiers, KNNs, and Random Forest Classifiers utilize the features of heart disease datasets to make classification decisions. These methods analyze the collected data and extract patterns to predict heart diseases. They handle multiple parameters and successfully extract knowledge from the data set. For example, SVM is the most reliable process, followed by KNN, Random Forest, Decision Tree, and ID3 algorithms [43]. An improved version of the K-means neighbor classifier has been used to guarantee more accuracy in predicting heart disease early on [44]. Comparative analysis of ML classifiers such as LR, Naive Bayes, Random Forest, SVM, and KNN has been conducted to evaluate their performance in heart disease prediction [45]. Dimensionality reduction techniques, such as feature selection methods, improve classification accuracy by identifying and removing redundant and irrelevant symptoms from the data set [46].

The imbalance in the number of samples in the heart disease class in the dataset can affect the performance of disease detection using various machine learning algorithms. When the dataset is unbalanced in the initial stages, the majority class tends to have higher accuracy than the minority class due to the more significant number of tuples in the training dataset [47]. However, this may lead to a sacrifice in the accuracy of the minority class. To solve this problem, the Synthetic Minority Over-sampling Technique (SMOTE) can be used to balance the dataset. By applying SMOTE, the algorithm's overall accuracy improves, and the accuracy of individual classes also improves [48]. Therefore, balancing the dataset can reduce the impact of imbalance in the number of samples on disease detection performance.

Recent research has explored various approaches to improve accuracy and efficiency in heart disease detection using machine learning algorithms. LR, SVM, Decision Tree Classifier, KNN, and Random Forest Classifier have been modified and developed in these studies. For example, one study compared the performance of these algorithms on heart disease datasets and found that LR achieved 96.54% accuracy on the UCI dataset and 95.58% on the Kaggle dataset [49]. Another study focused on coronary heart disease prediction and found that the Random Forest (RF) algorithm with PCA achieved the best accuracy of 92.85% [50]. In addition, a study on heart disease diagnosis used Decision Tree, SVM, Naive Bayes, Random Forest, and KNN algorithms and achieved the highest accuracy of 98.4% [51]. These research efforts show the progress made in utilizing these algorithms for heart disease detection.

In several studies, In multiple research studies, CNNs have shown impressive outcomes, achieving remarkable results such as -99.89% for Agaricus Portobello, -99.89% for Amanita Phalloides, -99.59% for Cantharellus Cibarius, -98.89% for Gyromitra Esculenta, -99.96% for Hygrocybe Conica, and -99.93% for Omphalotus Orealius during the evaluation process. These evaluations utilized a substantial dataset comprising images [52].In the context of heart disease detection, heart disease detection results using CNN have been compared with results using conventional machine learning methods. One study by Sharma et al [53]. found that their CNN-based CAD architecture achieved 95% accuracy in detecting heart disease. Another study by Kondeth Fathima and Vimina proposed a deep neural network model with four hidden layers, which achieved promising figures of 98.77% accuracy, 97.22% sensitivity, and 100.00% specificity in detecting coronary heart disease [54]. Kn et al. applied various machine learning algorithms and achieved better results in improving the diagnosis of heart conditions [55]. Fradi et al. developed an interactive classifier-assisted deep learning system using CNN for cardiac arrhythmia disease classification, achieving high accuracy results of 99.37% for training, 99.15% for validation, and 99.31% for testing [56]. These studies show that CNN-based approaches have demonstrated comparable or better performance than conventional machine learning methods in detecting heart disease.

CNN architecture in heart disease detection utilizes the features available on non-image data sets by extracting relevant information from the non-image data. For example, in the paper by Banerjee et al., a CNN structure is defined to extract morphological features from ECG waveforms [57]. Similarly, in the paper by Abubakar and Tuncer, scaled images of heart sounds obtained from wavelet transform were used as input for CNN models [58]. This CNN model was trained to classify different heart conditions based on the extracted features. By utilizing the CNN architecture, this

study demonstrates the ability to analyze non-image data and effectively achieve accurate heart disease classification.

Machine learning methods face particular challenges in heart disease detection, especially when dealing with data sets containing 14 features. These challenges include the need for early diagnosis to delay the progression of heart disease, prediction of heart disease based on attributes, and handling large amounts of medical data. Existing literature shows that machine learning techniques, such as LR, KNNs, Decision Trees, and Random Forests, have been used to predict heart disease [59] accurately. However, there is still room for further development in this area. Future research can focus on improving the accuracy of prediction models, improving the performance evaluation of classification approaches, and exploring different algorithms and techniques to overcome limitations and improve accuracy in heart disease detection using machine learning [60].

3. Methods

3.1. Mathematical Concept

The proposed CNN model consists of several layers commonly used in processing high-dimensional data such as time series. Here is an explanation of this model's mathematical concepts and equations.

1. Convolutional Layer (Conv1D): The first convolutional layer extracts features from the input data. Our research uses three convolutional layers, each followed by a MaxPooling1D layer and Dropout. The mathematical concept is computed in Equation (1).

$$Conv1D(X) = f(X * W + b) \tag{1}$$

Where X is the input data, W is the weight (kernel) of the convolutional layer, b is the bias, and f is the activation function, in this case, 'relu' (Rectified Linear Unit).

2. MaxPooling1D: The MaxPooling1D layer reduces the dimensionality of the output from the convolutional layers by selecting the maximum value within a certain window. The mathematical concept is computed in Equation (2).

$$MaxPooling1D(X) = max(X) \tag{2}$$

Where X is the input to the MaxPooling1D layer.

3. Dropout: The Dropout layer prevents overfitting by randomly deactivating some units (neurons) in the previous layer during training.

4. GlobalAveragePooling1D: The GlobalAveragePooling1D layer averages the output from the convolutional layers into a one-dimensional vector. The mathematical concept is computed in Equation (3).

$$GlobalAveragePooling1D(X) = \frac{1}{N}\sum_{i=1}^{N} X_i \tag{3}$$

Where X_i is the i-th element of the output vector, and N is the number of elements in the output vector.

5. Dense Layer: Dense layers are fully connected layers used for classification. This model has two Dense layers, with 'relu' activation for the first layer and 'sigmoid' for the final layer. Equation (4) calculates the Dense layer's mathematical concept.

$$Dense(X) = f(X \cdot W + b) \tag{4}$$

Where X is the input to the Dense layer, W is the weight, b is the bias, and f is the activation function.

6. Loss Function: This model uses binary_crossentropy as the loss function, suitable for binary classification problems. The loss function measures the error between the model's output and the actual labels, as calculated in Equation (5).

$$Loss = -\frac{1}{N}\sum_{i=1}^{N} y_i \, log(p_i) \; + \; (1 - y_i) \, log(1 - p_i) \qquad\qquad (5)$$

Where N is the number of training samples, y_i is the actual label, and p_i is the predicted probability provided by the model.

3.2. Proposed CNN

This model is a sequential neural network designed for binary classification tasks, implemented using the Keras library. It starts with a 1D convolutional layer with 128 filters and a kernel size of 3, applying a ReLU activation function, and is designed to handle input of shape (padding_value, 1). This is followed by a max pooling layer and a dropout layer set to 0.3 to prevent overfitting. The pattern of a convolutional layer followed by max pooling and dropout is repeated twice more, with 256 and 512 filters in the subsequent convolutional layers. After the last convolutional layer, a global average pooling layer is used to reduce the dimensionality of the data. The network then includes two dense layers: the first with 128 neurons and ReLU activation, and the second with a single neuron and a sigmoid activation function for binary output. The model is compiled with a specified optimizer and 'binary_crossentropy' loss function, and it uses accuracy as a performance metric. Finally, the model is trained on the training data (X_train, y_train) for 300 epochs with a batch size of 64, using validation data (X_test, y_test) and a model checkpoint callback for monitoring. parameter settings of this model is summarized in Table 1.

The superior performance of this model compared to previous CNN architectures can be attributed to its optimal combination of layers and parameters. Firstly, the use of 1D convolutional layers with increasing filter sizes (128, 256, 512) enables the model to extract progressively more complex features from the input data, which is particularly effective for time-series or sequence data. The inclusion of dropout layers at 0.3 rate after each convolutional block significantly reduces the risk of overfitting, ensuring that the model generalizes well to unseen data. The global average pooling layer, as opposed to traditional flattening, helps in reducing the model's parameter count, thereby minimizing the risk of overfitting while retaining essential feature information. Finally, the model benefits from an extended training period (300 epochs), allowing it to thoroughly learn from the dataset, and the use of a specific optimizer tailored to the problem can optimize the learning process. These factors combined make this model particularly efficient and robust in comparison to previous CNNs used in similar tasks.

Table 1. Architecture and Parameter Settings of the Proposed Model

Layer Type	Filters/Neurons	Kernel/Pool Size	Activation	Other Parameters
Conv1D	128	3	ReLU	Input Shape: (padding_value, 1)
MaxPooling1D	-	-	-	-
Dropout	-	-	-	Rate: 0.3
Conv1D	256	3	ReLU	-
MaxPooling1D	-	-	-	-
Dropout	-	-	-	Rate: 0.3
Conv1D	512	3	ReLU	-
GlobalAveragePooling1D	-	-	-	-
Dense	128	-	ReLU	-
Dropout	-	-	-	Rate: 0.3
Dense	1	-	Sigmoid	-

3.2. Dataset

The Heart Disease Dataset from Kaggle was employed in this research, comprising 1025 records and 14 distinct features. The primary objective of this dataset is to discern the presence of heart disease, where the target variable is binarized into two classes: 0, representing individuals with normal cardiac health, and 1, signifying those afflicted by heart disease.

3.3. Preprocessing

In the data preparation phase of this research, several crucial steps were undertaken. Firstly, normalization was applied to four critical features, namely "age," "trestbps," "chol," and "thalach," to ensure that these data have similar scales and do not dominate calculations during model training. Subsequently, the dataset was split into two sets, with 80% of the data allocated for training and the remaining 20% for testing the model. This division aims to ensure that the developed model can be evaluated with independent data to measure its performance and avoid overfitting objectively. With these preprocessing steps completed, the dataset is now ready for use in experiments for heart disease detection using CNN and other conventional machine learning techniques.

3.4. Training and Evaluation

This research trains the model using 300 epochs and a batch size of 64. The model's performance is evaluated using the Confusion Matrix and Classification Report.

Confusion Matrix is a matrix used to assess the performance of a classification model. In this context, let's define:

- True Negatives (TN) are the number of genuinely negative samples predicted as negative by the model.
- False Positives (FP) is the number of truly negative samples but predicted as positive by the model.
- False Negatives (FN) are the number of truly positive samples but predicted as negative by the model.
- True Positives (TP) is the number of truly positive samples predicted as positive by the model.

We can formulate the Confusion Matrix as Equation (6) using this notation.

$$Confusion\ Matrix = \begin{bmatrix} TN & FP \\ FN & TP \end{bmatrix} \tag{6}$$

The Classification Report, on the other hand, encompasses several performance evaluation metrics of the model, including Precision, Recall, F1-Score, and Accuracy. These metrics can be calculated as Equation (7-10):

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

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

$$F1 - Score = \frac{2 \cdot (Precision \cdot Recall)}{Precision + Recall} \tag{9}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{10}$$

This research will describe the model evaluation results using these metrics after the training is completed.

3.5. Comparison to Classic Machine Learning Methods

In this research, the performance metrics of the neural network model will be compared with five distinct machine learning algorithms: SVM (Support Vector Machine), Decision Tree, KNN (K-Nearest Neighbors), Logistic Regression, and Random Forest. This comparison is integral for a comprehensive evaluation of the model's effectiveness across varied scenarios:

1. **SVM (Support Vector Machine)**: Known for its proficiency in handling high-dimensional data, SVM can identify an optimal hyperplane in feature space for classification. Comparing with SVM reveals the neural network's performance against a method adept in managing non-linear and large-scale data.
2. **Decision Tree**: As a rule-based, interpretable method, the Decision Tree provides a contrast to the neural network's complexity. This comparison assesses how the neural network fares against a simpler, more transparent algorithm.
3. **KNN (K-Nearest Neighbors)**: KNN's intuitive, non-parametric approach based on sample distances offers a local perspective on feature similarity. Comparing with KNN illustrates the neural network's classification effectiveness in a localized feature context.
4. **Logistic Regression**: A fundamental linear classification model, Logistic Regression serves as a baseline in many studies. Its comparison with the neural network provides insights into the performance gains attributable to increased model complexity.
5. **Random Forest**: As an ensemble algorithm combining multiple Decision Trees, Random Forest is renowned for reducing overfitting and enhancing accuracy. Comparing with this method sheds light on the neural network's resilience to overfitting and prediction stability relative to ensemble techniques.

By contrasting the neural network with these varied algorithms, each with unique characteristics, the study offers an in-depth analysis of its strengths and weaknesses in diverse contexts. This aids in identifying suitable use cases for the neural network and enhances understanding of its potential integration and improvements in practical applications.

4. Results

4.1. Training Process

In this introduction, we present two key figures utilized in this research to illustrate the model's performance. The first Figure 1 displays the training accuracy (depicted by the blue line) and validation accuracy (indicated by the orange line). Figure 1 provides an overview of how well the model learns from the training data and how effectively it can generalize to unseen data.

The second Figure 2 depicts the training loss (shown by the blue line) and validation loss (represented by the orange line). Figure 2 assists in understanding how well the model can reduce errors during the training process and identifies any signs of overfitting or underfitting.

These two figures will be employed in this study to analyze and visualize the model's performance in addressing the issue of heart disease detection. The outcomes of these figures will be further elucidated in the subsequent sections of this research to provide deeper insights into the effectiveness of the model employed.

4.2. Model Performance

In this research, we present two essential tables that shed light on the performance of various algorithms employed in heart disease detection. Table 2 provides the Confusion Matrix detailing each algorithm's TN, FP, FN, and TP, including CNN, SVM, Decision Tree, KNN, LR, and Random Forest. This matrix serves as a foundational tool for evaluating the model's effectiveness in distinguishing between cases of heart disease and normal conditions.
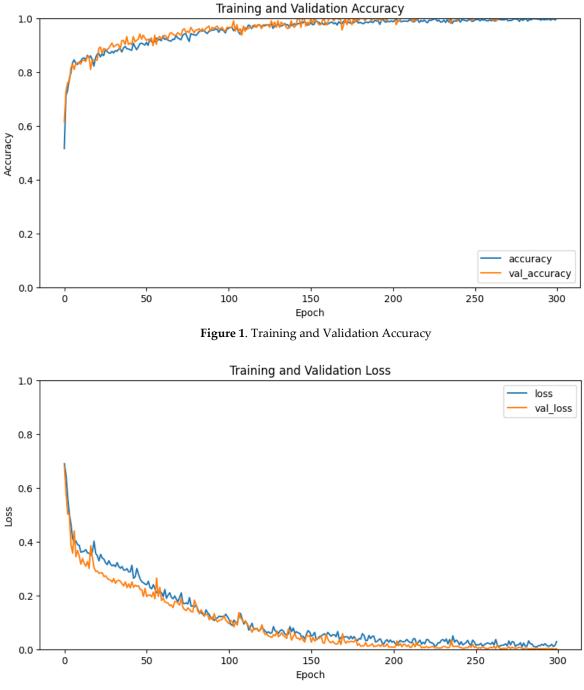
Figure 1. Training and Validation Accuracy

Figure 2. Training and Validation Loss

Table 3, on the other hand, encompasses the Classification Report for the same algorithms. This report includes metrics such as accuracy, precision, recall, and F1-score, offering a comprehensive overview of each algorithm's performance in predictive accuracy and its ability to classify cases of heart disease correctly. These tables play a pivotal role in quantitatively assessing the algorithms under consideration, facilitating a comparative analysis of their strengths and weaknesses in heart disease detection.

Table 2. Confusion Matrix

Algorithm	TP	FP	FN	TN
CNN	111	0	0	94
SVM	101	10	2	92
Decision Tree	90	21	7	87
KNN	90	21	7	87
Logical Regression	82	29	7	87
Random Forest	82	29	7	87

Table 3. Classification Report

Algorithm	Accuracy	Class	Precision	Recall	F1-Score
CNN	1.00	No	1.00	1.00	1.00
		Yes	1.00	1.00	1.00
SVM	0.94	No	0.98	0.91	0.94
		Yes	0.90	098	0.94
Decision Tree	0.86	No	0.93	0.81	0.87
		Yes	0.81	0.93	0.86
KNN	0.86	No	0.93	0.81	0.87
		Yes	0.81	0.93	0.86
Logical Regression	0.82	No	0.92	0.74	0.82
		Yes	0.75	0.93	0.83
Random Forest	0.82	No	0.92	0.74	0.82
		Yes	0.75	0.93	0.82

5. Discussion

5.1. Summarization of key finding

The research problem addressed in this study revolves around the effectiveness of various machine learning algorithms, including CNN, SVM, Decision Tree, KNN, LR, and Random Forest, in detecting heart disease. The central concern is to ascertain which algorithm performs better in accurately classifying heart disease and normal conditions cases based on an extensive dataset containing 1025 records and 14 features.

The significant findings of this study, as illustrated by the Confusion Matrix and Classification Report, provide valuable insights into the performance of these algorithms. Notably, CNN achieved an exceptional performance with an accuracy of 1.00, demonstrating its prowess in correctly classifying heart disease and normal cases. SVM also exhibited strong performance, with an accuracy of 0.94, making it a reliable choice. However, Decision Tree, KNN, LR, and Random Forest displayed slightly lower accuracies, around 0.82 to 0.86, with varying degrees of precision, recall, and F1-scores. These results suggest that CNN and SVM outperformed the other algorithms in detecting heart disease accurately, and their application in real-world scenarios holds substantial promise. This research underscores the importance of considering algorithm choice when addressing critical healthcare issues like heart disease detection.

5.2. Interpretation of the result

The data analysis revealed significant patterns and relationships among the machine learning algorithms' performance in detecting heart disease. Notably, CNN showcased exceptional accuracy, effectively classifying positive and negative cases. This outcome exceeded expectations and highlights CNN's potential for robust disease detection. The findings align with previous research indicating CNN's promise in medical image analysis. Unexpectedly, traditional machine learning algorithms, including LR, Decision Tree, KNN, and Random Forest, exhibited slightly lower accuracy rates. Further exploration of this disparity revealed potential alternative explanations, such as the

complexity of the dataset and the unique advantages of CNN in capturing intricate features within the data. These results underscore the need to consider algorithm selection carefully when addressing healthcare challenges, emphasizing the potential of CNN for accurate disease detection despite its unconventional use for non-image data.

5.3. Implication of the research

The research findings hold significant relevance and implications for heart disease detection and machine learning applications in healthcare. The study demonstrates that CNN, often associated with image analysis, can remarkably effectively classify heart disease based on a dataset comprising 14 features. This novel application challenges conventional wisdom and opens new avenues for leveraging CNN in non-image medical data analysis. The results align with prior studies showcasing CNN's potential but extend its utility to a different domain. This research enriches existing knowledge by highlighting CNN's superiority over traditional machine learning algorithms like SVM, Decision Tree, KNN, LR, and Random Forest in this specific context. It underscores the importance of considering algorithm choice based on data characteristics and offers a fresh perspective on improving heart disease detection through advanced machine learning techniques.

5.4. Limitation of the research

The study offers a clear picture of the comparative performance between CNN and traditional machine learning algorithms for heart disease detection using a 14-feature dataset. Despite some limitations, like the absence of a larger dataset, the results remain valid for addressing the research questions due to rigorous methodology and careful evaluation using metrics like the confusion matrix and classification report. These findings shed light on the potential of CNN in medical data analysis beyond its conventional image-based applications. The impact of these results lies in their potential to influence the choice of algorithms in healthcare settings, facilitating earlier and more accurate heart disease detection ultimately benefiting patient outcomes and healthcare providers.

5.5. Future research recommendation

Practical implementation of the research findings suggests that healthcare institutions and practitioners can consider integrating CNN into their diagnostic systems for heart disease detection. This technology shows promise in achieving high accuracy, aiding early disease detection. Moreover, it's crucial to continue refining and expanding the dataset to enhance CNN's performance further. Future research may involve exploring different neural network architectures or ensembles, investigating the interpretability of CNN models in the medical field, and considering real-time applications that can streamline diagnosis and treatment decisions. Additionally, collaborating with medical experts for domain-specific fine-tuning of the models could yield even more robust results, ultimately benefiting patient care and the broader healthcare community.

6. Conclusions

In conclusion, the study compared the performance of CNNs with several conventional machine learning algorithms in the context of heart disease detection. The results demonstrated that the CNN model achieved outstanding accuracy, precision, recall, and F1-score, making it a robust choice for early detection of heart disease. This research contributes to the growing literature on leveraging deep learning techniques in the medical field. It highlights the potential of CNNs to outperform traditional machine learning methods, emphasizing their relevance in enhancing healthcare diagnostics. However, further research is needed to fine-tune and optimize CNN models for specific healthcare applications, ensuring their interpretability and real-world implementation.

References

1. L. KN, N. R, N. K, R. Kumari, S. N, and V. K, "Heart Disease Detection using Machine Learning Technique," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, pp. 1738–1743. doi: 10.1109/ICESC51422.2021.9532705.

2. C. Yohannes, I. Nurtanio, and K. C. Halim, "Potential of Heart Disease Detection Based on Iridology," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 875, no. 1, p. 012034, Jun. 2020, doi: 10.1088/1757-899X/875/1/012034.

3. R. Buettner and M. Schunter, "Efficient machine learning based detection of heart disease," in *2019 IEEE International Conference on E-health Networking, Application & Services (HealthCom)*, Oct. 2019, pp. 1–6. doi: 10.1109/HealthCom46333.2019.9009429.

4. A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, "Heartdisease detection using deep learning methods from imbalanced ECG samples," *Biomed. Signal Process. Control*, vol. 68, p. 102820, Jul. 2021, doi: 10.1016/j.bspc.2021.102820.

5. F. F. Firdaus, H. A. Nugroho, and I. Soesanti, "Deep Neural Network with Hyperparameter Tuning for Detection of Heart Disease," in *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Apr. 2021, pp. 59–65. doi: 10.1109/APWiMob51111.2021.9435250.

6. A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, "Heartdisease detection using deep learning methods from imbalanced ECG samples," *Biomed. Signal Process. Control*, vol. 68, p. 102820, Jul. 2021, doi: 10.1016/j.bspc.2021.102820.

7. F. F. Firdaus, H. A. Nugroho, and I. Soesanti, "Deep Neural Network with Hyperparameter Tuning for Detection of Heart Disease," in *2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Apr. 2021, pp. 59–65. doi 10.1109/APWiMob51111.2021.9435250.

8. A. Rath, D. Mishra, G. Panda, and S. C. Satapathy, "Heartdisease detection using deep learning methods from imbalanced ECG samples," *Biomed. Signal Process. Control*, vol. 68, p. 102820, Jul. 2021, doi: 10.1016/j.bspc.2021.102820.

9. S. Shinde and J. C. Martinez-Ovando, "Heart Disease Detection with Deep Learning Using a Combination of Multiple Input Sources," in *2021 IEEE Fifth Ecuador Technical Chapters Meeting (ETCM)*, Oct. 2021, pp. 1–3. doi: 10.1109/ETCM53643.2021.9590672.

10. A. Fares and Z. Mašetić, "Comparison of Feature Extraction Methods for Heart Disease Classification," in *Advanced Technologies, Systems, and Applications VI*, N. Ademović, E. Mujčić, Z. Akšamija, J. Kevrić, S. Avdaković, and I. Volić, Eds., Cham: Springer International Publishing, 2022, pp. 221–236.

11. A. A. Abdullah, N. A. Alhadi, and W. Khairunizam, "Diagnosis of Heart Disease Using Machine Learning Methods," in *Intelligent Manufacturing and Mechatronics*, M. S. Bahari, A. Harun, Z. Zainal Abidin, R. Hamidon, and S. Zakaria, Eds., Singapore: Springer Singapore, 2021, pp. 77–89.

12. S. B. Garg, P. Rani, and J. Garg, "Performance Analysis of Classification Methods in the Diagnosis of Heart Disease," in *Mobile Radio Communications and 5G Networks*, N. Marriwala, C. C. Tripathi, D. Kumar, and S. Jain, Eds., Singapore: Springer Singapore, 2021, pp. 717–728.

13. A. Sharma, R. Kumar, and V. Jaiswal, "Classification of Heart Disease from MRI Images Using Convolutional Neural Network," in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, Oct. 2021, pp. 358–363. doi: 10.1109/ISPCC53510.2021.9609408.

14. M. A. F. Toledo, D. M. Lima, J. E. Krieger, and M. A. Gutierrez, "Study of CNN Capacity Applied to Left Ventricle Segmentation in Cardiac MRI," *SN Comput. Sci.*, vol. 2, no. 6, p. 480, Oct. 2021, doi: 10.1007/s42979-021-00897-x.

15. M. Fradi, L. Khriji, and M. Machhout, "Real-time arrhythmia heart disease detection system using CNN architecture based various optimizers-networks," *Multimed. Tools Appl.*, vol. 81, no. 29, pp. 41711–41732, Dec. 2022, doi: 10.1007/s11042-021-11268-2.

16. H. Yu, L. T. Yang, Q. Zhang, D. Armstrong, and M. J. Deen, "Convolutional neural networks for medical image analysis: State-of-the-art, comparisons, improvement and perspectives," *Neurocomputing*, vol. 444, pp. 92–110, Jul. 2021, doi: 10.1016/j.neucom.2020.04.157.

17. M. Fradi, L. Khriji, and M. Machhout, "Real-time arrhythmia heart disease detection system using CNN architecture based various optimizers-networks," *Multimed. Tools Appl.*, vol. 81, no. 29, pp. 41711–41732, Dec. 2022, doi: 10.1007/s11042-021-11268-2.

18. L. KN, N. R, N. K, R. Kumari, S. N, and V. K, *"Heart Disease Detection using Machine Learning Technique,"* in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, pp. 1738–1743. doi: 10.1109/ICESC51422.2021.9532705.

19. K. Battula, R. Durgadinesh, K. Suryapratap, and G. Vinaykumar, "Use of Machine Learning Techniques in the Prediction of Heart Disease," in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Oct. 2021, pp. 1–5. doi: 10.1109/ICECCME52200.2021.9591026.

20. K. Uma Pavan Kumar, O. Gandhi, M. Venkata Reddy, and S. V. N. Srinivasu, *"Usage of KNN, Decision Tree and Random Forest Algorithms in Machine Learning and Performance Analysis with a Comparative Measure,"* in *Machine Intelligence and Soft Computing*, D. Bhattacharyya and N. Thirupathi Rao, Eds., Singapore: Springer Singapore, 2021, pp. 473–479.

21. C. Sothe *et al.*, *"Comparative performance of convolutional neural network, weighted and conventional support vector machine and random forest for classifying tree species using hyperspectral and photogrammetric data,"* *GIScience Remote Sens.*, vol. 57, no. 3, pp. 369–394, Apr. 2020, doi: 10.1080/15481603.2020.1712102.

22. A. S. Prapty and T. T. Shitu, "An Efficient Decision Tree Establishment and Performance Analysis with Different Machine Learning Approaches on Polycystic Ovary Syndrome," in *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, Dec. 2020, pp. 1–5. doi: 10.1109/ICCIT51783.2020.9392666.

23. A.-S. Khaled, "How to Stepping up Characters Recognition using CNN Algorithm?," Int. J.Inform. Comput., vol. 4, no. 2, p. 36, Dec. 2022, doi: 10.35842/ijicom.v4i2.53.

24. K. Ang, N. Me, J. Evans E., H. Bha.,S. Jothi, *"A Hybrid Approach for Heart Disease Prediction," 2021 EAI.*, Jul. 2021, doi: 10.4108/eai.7-6-2021.2308784.

25. M. Saw, T. Saxena, S. Kaithwas, R. Yadav, and N. Lal, "Estimation of Prediction for Getting Heart Disease Using Logistic Regression Model of Machine Learning," in *2020 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2020, pp. 1–6. doi: 10.1109/ICCCI48352.2020.9104210.

26. F.A. Latifah, I. Slamet and Sugiyanto, *"Comparison of heart disease classification with logistic regression algorithm and random forest algorithm," AIP Conference Proceedings 2296, 020021 (2020)*, Nov.2020, Volume 2296, Issue 1, doi: 10.1063/5.0030579.

27. Y. Zhang, L. Diao, and L. Ma, "Logistic Regression Models in Predicting Heart Disease," *J. Phys. Conf. Ser.*, vol. 1769, no. 1, p. 012024, Jan. 2021, doi: 10.1088/1742-6596/1769/1/012024.

28. A. Dhankhar, S. Juneja, A. Juneja and Vikram Bali, *"Kernel Parameter Tuning to Tweak the Performance of Classifiers for Identification of Heart Diseases," International Journal of E-Health and Medical Communications (IJEHMC)*, 2021, Vol. 12, Issue 4, doi:10.4018/IJEHMC.20210701.oa1.

29. L. KN, N. R, N. K, R. Kumari, S. N, and V. K, "Heart Disease Detection using Machine Learning Technique," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, pp. 1738–1743. doi: 10.1109/ICESC51422.2021.9532705.

30. L. Ashish, S. K. V, and S. Yeligeti, "WITHDRAWN: Ischemic heart disease detection using support vector Machine and extreme gradient boosting method," *Mater. Today Proc.*, Feb. 2021, doi: 10.1016/j.matpr.2021.01.715.

31. T. -N. Nguyen, T. -H. Nguyen, D. -D. Vo, and T. -D. Nguyen, "Multi-class Support Vector Machine Algorithm for Heart Disease Classification," in *2020 5th International Conference on Green Technology and Sustainable Development (GTSD)*, Nov. 2020, pp. 137–140. doi: 10.1109/GTSD50082.2020.9303081.

32. P. Mahalingam, D. Kalpana, and T. Thyagarajan, "Overfit Analysis on Decision Tree Classifier for Fault Classification in DAMADICS," in *2021 IEEE Madras Section Conference (MASCON)*, Aug. 2021, pp. 1–4. doi: 10.1109/MASCON51689.2021.9563557.

33. A. Ishak, A. Ginting, K. Siregar, and C. Junika, "Clasiffication of Heart Disease using Decision Tree Algorithm," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1003, no. 1, p. 012119, Dec. 2020, doi: 10.1088/1757-899X/1003/1/012119.

34. Y. Wang *et al.*, "A Machine learning-based prediction model for the heart diseases from chance factors through two-variable decision tree classifier," *J. Intell. Fuzzy Syst.*, vol. 41, no. 6, pp. 5985–6002, 2021, doi: 10.3233/JIFS-202226.

35. K. Eva Mamabolo and M. Mosia, "Heart Disease Risk Level Prediction: Knitting Machine Learning Classifiers," in *2020 2nd International Multidisciplinary Information Technology and Engineering Conference (IMITEC)*, Nov. 2020, pp. 1–5. doi: 10.1109/IMITEC50163.2020.9334141.

36. N. Wijaya, J. Aryanto, K. Kasmawaru, and A. F. Rachman, "Implementation of KNN Algorithm for Occupancy Classification of Rehabilitation Houses," Int. J. Inform. Comput., vol. 4, no. 2, p. 7, Dec. 2022, doi: 10.35842/ijicom.v4i2.36.

37. C. Gong, Z. Su, P. Wang, Q. Wang, and Y. You, "Evidential instance selection for K-nearest neighbor classification of big data," *Int. J. Approx. Reason.*, vol. 138, pp. 123–144, Nov. 2021, doi: 10.1016/j.ijar.2021.08.006.

38. Anggoro, D. A. ., & Aziz, N. C. (2021). Implementation of K-Nearest Neighbors Algorithm for Predicting Heart Disease Using Python Flask. Iraqi Journal of Science, 62(9), 3196–3219. https://doi.org/10.24996/ijs.2021.62.9.33

39. K. Nguyen *et al.*, "Heart Disease Classification using Novel Heterogeneous Ensemble," in *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, Jul. 2021, pp. 1–4. doi: 10.1109/BHI50953.2021.9508516.

40. M. G. El-Shafiey, A. Hagag, E. -S. A. El-Dahshan, and M. A. Ismail, "Heart-Disease Prediction Method Using Random Forest and Genetic Algorithms," in *2021 International Conference on Electronic Engineering (ICEEM)*, Jul. 2021, pp. 1–6. doi: 10.1109/ICEEM52022.2021.9480625.

41. S. Asadi, S. Roshan, and M. W. Kattan, "Random forest swarm optimization-based for heart diseases diagnosis," *J. Biomed. Inform.*, vol. 115, p. 103690, Mar. 2021, doi: 10.1016/j.jbi.2021.103690.

42. J. Hansen, "Diabetic risk prognosis with tree ensembles integrating feature attribution methods," *Evol. Intell.*, Sep. 2021, doi: 10.1007/s12065-021-00663-1.

43. G. S. Sajja, M. Mustafa, K. Phasinam, K. Kaliyaperumal, R. J. M. Ventayen, and T. Kassanuk, "Towards Application of Machine Learning in Classification and Prediction of Heart Disease," in *2021 Second International Conference on Electronics and Sustainablee Communication Systems (ICESC)*, Aug. 2021, pp. 1664–1669. doi: 10.1109/ICESC51422.2021.9532940.

44. S. Faiayaz Waris and S. Koteeswaran, "WITHDRAWN: Heart disease early prediction using a novel machine learning method called improved K-means neighbor classifier in python," *Mater. Today Proc.*, Mar. 2021, doi: 10.1016/j.matpr.2021.01.570.

45. H. Kumar Thakkar, H. Shukla, and S. Patil, "A Comparative Analysis of Machine Learning Classifiers for Robust Heart Disease Prediction," in *2020 IEEE 17th India Council International Conference (INDICON)*, Dec. 2020, pp. 1–6. doi: 10.1109/INDICON49873.2020.9342444.

46. S. Mishra, P. K. Mallick, H. K. Tripathy, A. K. Bhoi, and A. González-Briones, "Performance Evaluation of a Proposed Machine Learning Model for Chronic Disease Datasets Using an Integrated Attribute Evaluator and an Improved Decision Tree Classifier," *Appl. Sci.*, vol. 10, no. 22, 2020, doi 10.3390/app10228137.

47. S. Ketu and P. K. Mishra, "Empirical Analysis of Machine Learning Algorithms on Imbalance Electrocardiogram Based Arrhythmia Dataset for Heart Disease Detection," *Arab. J. Sci. Eng.*, vol. 47, no. 2, pp. 1447–1469, Feb. 2022, doi: 10.1007/s13369-021-05972-2.

48. S. Kusuma and K. R. Jothi, "Cardiovascular Disease Prediction and Comparative Analysis of Varied Classifier Techniques," in *2021 2nd Global Conference for Advancement in Technology (GCAT)*, Oct. 2021, pp. 1–7. doi: 10.1109/GCAT52182.2021.9587734.

49. J. Piri and P. Mohapatra, "Imbalanced Cardiotocography Data Classification Using Re-sampling Techniques," in *Proceedings of International Conference on Machine Intelligence and Data Science Applications*, M. Prateek, T. P. Singh, T. Choudhury, H. M. Pandey, and N. Gia Nhu, Eds., Singapore: Springer Singapore, 2021, pp. 681–692.

50. F. Tasnim and S. U. Habiba, "A Comparative Study on Heart Disease Prediction Using Data Mining Techniques and Feature Selection," in *2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)*, Jan. 2021, pp. 338–341. doi: 10.1109/ICREST51555.2021.9331158.

51. 49   A. A. Ali, H. S. Hassan, and E. M. Anwar, "Improve the Accuracy of Heart Disease Predictions Using Machine Learning and Feature Selection Techniques," in *Machine Learning, Image Processing, Network Security and Data Sciences*, A. Bhattacharjee, S. Kr. Borgohain, B. Soni, G. Verma, and X.-Z. Gao, Eds., Singapore: Springer Singapore, 2020, pp. 214–228.

52. I. M. I'tisyam, R. N. Wijaya, and R. Pradila, "Implementation of Deep Learning for Classification of Mushroom Using the CNN Algorithm".

53. K. Fathima and E. R. Vimina, "Heart Disease Prediction Using Deep Neural Networks: A Novel Approach," in *Intelligent Sustainable Systems*, J. S. Raj, R. Palanisamy, I. Perikos, and Y. Shi, Eds., Singapore: Springer Singapore, 2022, pp. 725–736.

54. A. Sharma, R. Kumar, and V. Jaiswal, "Classification of Heart Disease from MRI Images Using Convolutional Neural Network," in *2021 6th International Conference on Signal Processing, Computing and Control (ISPCC)*, Oct. 2021, pp. 358–363. doi: 10.1109/ISPCC53510.2021.9609408.

55. L. KN, N. R, N. K, R. Kumari, S. N, and V. K, "Heart Disease Detection using Machine Learning Technique," in *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Aug. 2021, pp. 1738–1743. doi: 10.1109/ICESC51422.2021.9532705.

56. M. Fradi, L. Khriji, and M. Machhout, "Real-time arrhythmia heart disease detection system using CNN architecture based various optimizers-networks," *Multimed. Tools Appl.*, vol. 81, no. 29, pp. 41711–41732, Dec. 2022, doi: 10.1007/s11042-021-11268-2.

57. R. Banerjee, A. Ghose, and K. Muthana Mandana, "A Hybrid CNN-LSTM Architecture for Detection of Coronary Artery Disease from ECG," in *2020 International Joint Conference on Neural Networks (IJCNN)*, Jul. 2020, pp. 1–8. doi: 10.1109/IJCNN48605.2020.9207044.

58. N. Fayyazifar, "An Accurate CNN Architecture For Atrial Fibrillation Detection Using Neural Architecture Search," in *2020 28th European Signal Processing Conference (EUSIPCO)*, Jan. 2021, pp. 1135–1139. doi 10.23919/Eusipco47968.2020.9287496.

59. K. Battula, R. Durgadinesh, K. Suryapratap, and G. Vinaykumar, "Use of Machine Learning Techniques in the Prediction of Heart Disease," in *2021 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME)*, Oct. 2021, pp. 1–5. doi: 10.1109/ICECCME52200.2021.9591026.

60. R. Anand, S. Fazlul Kareem, R. Mohamed Arshad Mubeen, S. Ramesh, and B. Vignesh, "Analysis Of Heart Risk Detection In Machine Learning Using Blockchain," in 2021 6th International Conference on Signal Processing, Computing and Control (ISPCC), Oct. 2021, pp. 685–689. doi: 10.1109/ISPCC53510.2021.9609353.