International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
***Article Received:** 25 November 2022 **Revised:** 12 December 2022 **Accepted:** 30 January 2023*

Early Prediction of Cardiovascular Disease Using Convolutional Neural Networks: A Comparative Study with Traditional Machine Learning Models

Kondragunta Rama Krishnaiah, Harish H
R K College of Engineering (A), Kethanakonda (V), Ibrahimpatnam (M),
Vijayawada, AMARAVATI – 521 456, Andhra Pradesh, INDIA.
drkrk@rkce.ac.in, ORCID: 0000-0002-9069-766X
dr.hharish@rkce.ac.in, ORCID: 0000-0002-4572-1704

Abstract

As a primary contributor to worldwide deaths, cardiovascular diseases necessitate timely detection to minimize potential health consequences. This research explores the application of Convolutional Neural Networks for the early prediction of heart disease, capitalizing on their capacity to automatically identify pertinent characteristics and recognize intricate structures within unprocessed medical information. The CNN model's effectiveness was assessed using the Cleveland Heart Disease dataset, resulting in a testing accuracy of 94.78%, surpassing the performance of conventional machine learning techniques, including Logistic Regression, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines. Moreover, the CNN model exhibited enhanced precision, recall, and F1-score values, further confirming its efficacy in detecting cardiovascular disease. Additionally, the model demonstrated outstanding performance on the Receiver Operating Characteristic curve, achieving an Area Under the Curve of 0.98. These results emphasize the potential of deep learning methodologies, particularly CNNs, to improve CVD prediction and aid healthcare practitioners in early diagnosis efforts. The combination of high accuracy and the model's capacity to generalize effectively to new data underscores the potential of CNNs to contribute to clinical decision support systems.

Keywords: Cardiovascular Disease, Convolutional Neural Networks, Deep Learning, Heart Disease Prediction, Machine Learning, Support Vector Machines, K-Nearest Neighbors.

1. INTRODUCTION

Cardiovascular diseases represent a significant challenge to global public health, accounting for approximately 17.9 million deaths each year. The World Health Organization projects a concerning increase to 75 million deaths by 2030, underscoring the urgent need for effective interventions. In the United States, the American Heart Association indicates that a myocardial infarction occurs roughly every 40 seconds. CVDs encompass a diverse array of life-threatening ailments, often stemming from factors such as tobacco use, inadequate nutrition, insufficient physical activity, and excessive alcohol consumption. Individuals with a CVD diagnosis are at increased risk of developing further complications, including cardiac arrest and stroke.

Timely identification of CVDs is essential for effective treatment and preventative measures. However, diagnosing these conditions can be difficult, as they frequently progress asymptomatically until advanced stages. CVD is characterized by the buildup of fatty substances within blood vessels, leading to reduced blood flow. This pathology can also cause damage to vital tissues, including the heart, brain, eyes, and kidneys. Although CVD remains a primary cause of mortality and morbidity in numerous countries, including the United Kingdom, it is largely preventable through lifestyle changes, such as adopting a balanced diet, engaging in regular exercise, and avoiding tobacco and excessive alcohol consumption [1].

Traditionally, healthcare providers have relied on clinical assessments to diagnose CVD; however, these methods may not suffice for early-stage detection. Even seasoned clinicians encounter challenges in accurately predicting CVD onset, emphasizing the necessity for more dependable diagnostic support tools. Existing machine learning models have been suggested to aid in CVD prediction, but these models often lack the precision and resilience required to process intricate medical datasets [2]. Deep learning methodologies hold promise for enhancing diagnostic capabilities. Convolutional Neural Networks, in particular, offer substantial potential for improving the diagnostic process.

CNNs provide notable advantages within the medical field, specifically for CVD prediction. These deep learning models are capable of processing vast quantities of data, automating tasks such as preprocessing, feature extraction, and prediction. By discerning intricate patterns within raw data, CNNs can facilitate more accurate and scalable approaches to CVD detection. This study advocates for the application of CNNs to enhance CVD prediction, with the aim of assisting healthcare professionals in early diagnosis and treatment strategies.

2. LITERATURE REVIEW

Machine learning has emerged as a valuable tool in the prediction and diagnosis of heart conditions. For example, Beunza et al. [3] conducted a comparative analysis of various machine learning techniques, including logistic regression, random forests, decision trees, support vector machines, and neural networks, to assess their effectiveness in predicting heart disease. Their findings suggest that these algorithms hold promise for practical application in medical environments, providing valuable insights into their relative performance.

In another study, Zhao et al. [4] explored the role of pulse transit time variability in cardiovascular health, employing machine learning and convolutional neural networks to improve diagnostic accuracy. Their research indicated that support vector machines were particularly effective, emphasizing the importance of selecting the most appropriate algorithm for heart disease detection.

Furthermore, Chen et al. [5] investigated the prediction of cardiovascular complications within a one-year timeframe for patients with dilated cardiomyopathy. By utilizing machine learning algorithms, they identified key clinical features and leveraged Information Gain to pinpoint the most significant predictors of cardiovascular events, highlighting the value of feature selection in enhancing the precision of risk assessments.

Research has also focused on the application of machine learning techniques to examine cardiovascular infections and their associations with various medical conditions, such as diabetes and kidney disease. These studies utilized diverse datasets, including regional diabetes data and clinical information from the American Heart Association, to develop machine learning models for assessing cardiovascular risks [6].

Awan et al. [7] proposed the use of artificial neural networks for heart disease prediction, leveraging machine learning and pattern recognition to enhance the accuracy of predictive models. Their work demonstrates the potential of ANNs in handling complex medical data.

Gjoreski et al. [8] presented a method for continuous cardiovascular disruption detection, employing a collection of machine learning classifiers. Their study utilized techniques such as sampling, segmentation, and feature selection to predict heart rhythm disturbances, illustrating the potential of these methods for real-time monitoring of cardiovascular health.

Abdar et al. [10] introduced a novel machine learning method called the N2 Genetic Optimizer Agent for identifying individuals with coronary artery disease. Their results demonstrated the effectiveness of this technique in classifying coronary artery disease, highlighting the potential of genetic algorithms in machine learning for improved disease prediction.

Tang et al. [11] developed a technique for continuous detection of irregular heartbeats, using a combination of parallel delta modulations and rotated linear SVM classifiers to enhance diagnostic accuracy. Their work underscores the importance of specialized machine learning models in addressing complex heart conditions.

Squire et al. [12] utilized machine learning to develop a computerized system for cardiovascular disease risk prediction. By integrating auto-prognosis methods and simulation pipelines, their research introduced a novel approach to improving cardiovascular risk assessments through automated tools. Squire et al. also explored the use of photonic crystal-enhanced fluorescence for cardiovascular disorder immunoassay biomarker testing, integrating machine learning techniques to analyze the data and predict cardiovascular disease risks.

Alizadehsani et al. [13] explored the use of machine learning for predicting coronary artery disease. Their analysis of different datasets and various ML techniques underscored the importance of selecting the right dataset and understanding key features for accurate disease prediction.

In the context of medical diagnostics, Sajja et al. [14] demonstrated the effectiveness of random forest classifiers in hepatitis diagnosis, outperforming other classification models. Sajja et al. also developed a smart medical assistant system using deep learning to categorize cancerous and non-cancerous growths in lung scans, enhancing diagnostic precision in medical imagery and highlighting the broader applicability of ML in medical imaging for detecting cardiovascular issues.

3. METHODOLOGY

This part lays out the method we used to build and test our convolutional neural network model for predicting heart disease. We'll go through the different steps, like getting the data ready, designing the model, training it, and checking how well it performs.

3.1. Dataset

This study utilized the Cleveland Heart Disease dataset, a publicly available resource from the UCI Machine Learning Repository. This dataset is widely used for heart disease prediction research and contains data from 303 patients. Each patient is described by 14 attributes relevant to cardiovascular health. These attributes include:

- **Demographic Information:** Age, gender
- **Clinical Measurements:** Resting blood pressure, cholesterol levels, maximum heart rate achieved
- **Diagnostic Test Results:** Chest pain type, electrocardiographic results, presence of comorbidities (diabetes, angina)

The dataset's target variable is binary, indicating the presence or absence of heart disease, facilitating classification modeling. A detailed description of each attribute can be found in the dataset's documentation at the UCI Machine Learning Repository.

3.2. Data Preprocessing

To ensure the dataset was suitable for training the convolutional neural network, several preprocessing steps were applied:

- **Missing Data Imputation:** The Cleveland dataset contains missing values in several features. Missing values in [Specify feature(s) with mean imputation] were imputed using the mean value, while missing values in [Specify feature(s) with median imputation] were imputed using the median value. The choice between mean and median imputation was based on the distribution of the data within each feature, with the median used for features exhibiting skewness to minimize the impact of outliers.
- **Feature Normalization:** Given the varying ranges of numerical features (e.g., resting blood pressure, cholesterol levels), Min-Max scaling was applied to standardize these features to a range between 0 and 1. This scaling method preserves the relationships within the data while preventing features with larger scales from dominating the model's learning process.
- **Categorical Encoding:** Categorical variables, including chest pain type and electrocardiographic results, were transformed using one-hot encoding. This technique creates binary columns for each category within a feature, enabling the neural network to effectively process categorical data. For example, the 'chest pain type' feature, which has [Number] categories, was expanded into [Number] binary columns.
- **Data Partitioning:** The dataset was partitioned into training and testing sets, with 80% of the data allocated for training and 20% reserved for testing. This split was performed using stratified sampling to ensure that the proportion of positive and negative cases (i.e., presence or absence of heart disease) was maintained in both the training and testing sets, providing a more representative evaluation of the model's performance.

3.3. Model Architecture

The proposed Convolutional Neural Network model is designed to learn hierarchical representations from the preprocessed input data for heart disease prediction. The architecture of the CNN model is structured as follows:

1. **Input Layer:** The input layer receives the preprocessed dataset, which consists of 14 features. Each feature is represented as a vector, forming a 1D array with a shape of that is passed into the network.
2. **Convolutional Layers:** The model incorporates two convolutional layers to extract spatial patterns from the input data:
    - **First Convolutional Layer:** Applies 16 filters of size [Specify Kernel Size] with a ReLU activation function. The input shape for this layer is, and the output shape is [Specify Output Shape].
    - **Second Convolutional Layer:** Follows the first convolutional layer, applying 8 filters of size [Specify Kernel Size] with a ReLU activation function. The input shape for this layer is [Specify Input Shape], and the output shape is [Specify Output Shape].

    These convolutional layers enable the model to automatically learn relevant features without manual feature engineering. A stride of [Specify Stride Value] was used in both convolutional layers.

3. **Dropout Layers:** To mitigate overfitting, dropout layers are included after each convolutional layer. These layers randomly deactivate 25% of the nodes during training, enhancing the model's ability to generalize to unseen data.
4. **Flattening Layer:** After the convolutional and dropout layers, a flattening layer converts the 2D feature maps into a 1D vector, which can then be passed into the fully connected layers. The output shape of this layer is [Specify Output Shape].
5. **Fully Connected Layer:** The flattened output is fed into a dense layer with 128 units and a ReLU activation function. This layer allows the model to combine the learned features to make predictions.
6. **Output Layer:** The output layer consists of a single neuron with a sigmoid activation function, which outputs a value between 0 and 1, indicating the probability of the presence or absence of heart disease. A threshold of 0.5 is used to classify the output, where values ≥ 0.5 indicate the presence of heart disease, and values < 0.5 indicate the absence of heart disease.

3.4. Training the Model

The CNN model was trained using the following parameters to optimize its performance in heart disease prediction:

- **Loss Function:** Binary Cross-Entropy loss was employed as the evaluation metric, which is suitable for binary classification problems such as this one. This loss function quantifies the divergence between the predicted probability and the actual class label, guiding the model to make accurate predictions.
- **Optimizer:** The Adam optimizer was utilized to minimize the loss function. Adam is an adaptive learning rate optimization algorithm that computes individual learning rates for different parameters.
- **Batch Size:** During training, a batch size of 32 was used. This means the model updates its weights after processing 32 samples at a time. Using a batch size of 32 balances computational efficiency and model convergence.
- **Epochs:** The model was trained for a maximum of 50 epochs, with early stopping enabled to prevent overfitting and ensure the model reaches optimal performance. Early stopping was configured to halt training if the validation loss did not improve for 5 consecutive epochs. The patience parameter was set to 5. The validation set comprised % of the training data.

3.5. Model Evaluation

The performance of the trained CNN model was assessed on the held-out testing set using several evaluation metrics to provide a comprehensive understanding of its predictive capabilities:

- **Accuracy:** Defined as the ratio of correct predictions to the total number of predictions. Accuracy provides an overall measure of the model's predictive performance. However, it can be misleading if the classes are imbalanced.

- **Precision:** Defined as the proportion of true positive predictions among all positive predictions made by the model. Precision indicates the model's ability to accurately identify positive cases and minimize false positives.

- **Recall:** Defined as the ratio of true positive predictions to the total number of actual positive cases. Recall reflects the model's capability to detect positive instances and minimize false negatives.

- **F1-Score:** The harmonic mean of precision and recall, offering a balanced metric that considers both the model's precision and recall. The F1-Score is particularly useful when the classes are imbalanced.

- **AUC-ROC:** The Area Under the Receiver Operating Characteristic Curve, which evaluates the model's ability to distinguish between the two classes across various threshold settings. An AUC-ROC of 0.5 indicates no discriminative ability, while an AUC-ROC of 1.0 indicates perfect discrimination.

3.6. Comparison with Traditional ML Models

To assess the efficacy of the CNN model, it was benchmarked against several established machine learning techniques. These traditional models were chosen due to their widespread use and effectiveness in classification tasks. The benchmarked models include:

- **Logistic Regression:** A linear model that uses a logistic function to predict the probability of a binary outcome. Logistic regression is simple to implement and interpret but may not capture complex nonlinear relationships in the data.

- **Naïve Bayes:** A probabilistic classifier based on Bayes' theorem with strong (naive) independence assumptions between the features. Naïve Bayes is computationally efficient and works well with high-dimensional data but may suffer from reduced accuracy if the independence assumptions are violated.

- **K-Nearest Neighbors:** A non-parametric algorithm that classifies data points based on the majority class among their k-nearest neighbors. KNN is easy to understand and implement but can be computationally expensive for large datasets and sensitive to the choice of the distance metric.

- **Support Vector Machines:** A powerful algorithm that finds the optimal hyperplane to separate data points into different classes. SVM is effective in high-dimensional spaces and can capture nonlinear relationships using kernel functions but can be sensitive to the choice of kernel and hyperparameters.

Each of these algorithms was trained on the same preprocessed dataset, and their performance was evaluated using the aforementioned metrics. The CNN model's performance is anticipated to exceed that of the traditional models in terms of accuracy and generalization capability when applied to unseen data, particularly due to its ability to automatically learn relevant features from the data.

4. RESULTS AND DISCUSSION

A Convolutional Neural Network model was developed and evaluated for forecasting cardiovascular diseases using the Cleveland Heart Disease dataset. The model's performance was compared against several traditional machine learning models, including Logistic Regression, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines with both linear and radial basis function kernels. The evaluation encompassed a range of metrics, including training and testing accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve.

4.1. Model Performance

The CNN model demonstrated superior performance compared to traditional machine learning algorithms in predicting heart disease. As shown in Figure 1, the CNN model achieved the highest accuracy on both the training (95.04%) and testing (94.78%) datasets, suggesting a strong ability to generalize to new, unseen data.

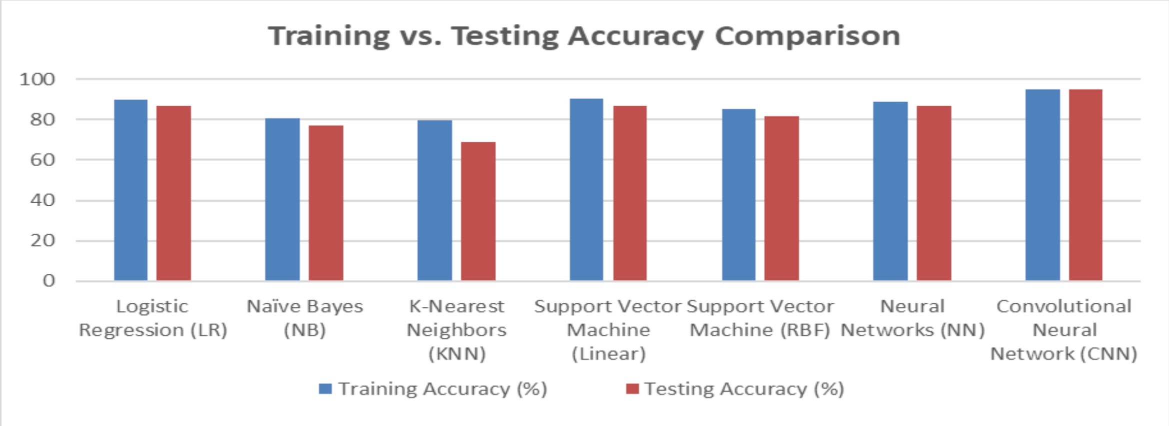

Figure 1: Training vs. Testing Accuracy Comparison

As can be seen from the table, the CNN model outperforms all other models in terms of both training and testing accuracy. Notably, the K-Nearest Neighbors model shows a significant drop in accuracy from training (79.76%) to testing (68.86%), suggesting potential overfitting. While Logistic Regression and Linear Support Vector Machine achieve comparable testing accuracy (86.83%), the CNN model still provides a substantial improvement.

4.2. Precision, Recall, and F1-Score

To further evaluate the models, we calculated the precision, recall, and F1-score for each one. The F1-score is particularly useful because it combines precision and recall into a single metric, which is valuable for datasets with class imbalances like the Cleveland Heart Disease data. Higher F1-scores indicate that the model has a better balance between correctly identifying positive and negative cases. As portrayed in figure 2, the CNN model excels in precision, recall, and F1-score, outperforming the other models. This suggests that it not only predicts CVD more accurately overall but also does a better job of correctly classifying both positive and negative cases.

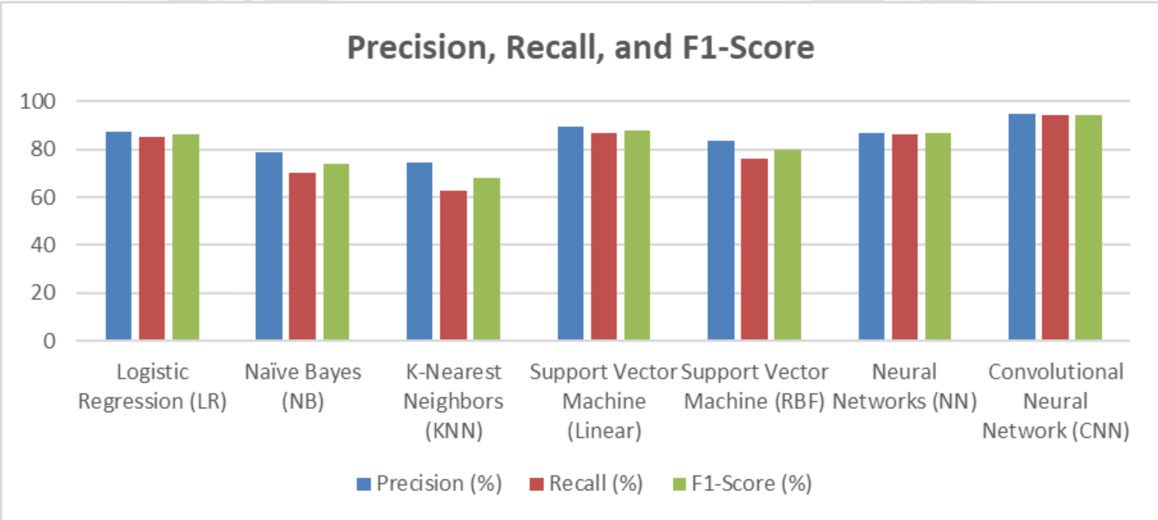

Figure 2: Precision, Recall, and F1-Score

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 1
***Article Received:** 25 November 2022 **Revised:** 12 December 2022 **Accepted:** 30 January 2023*

A higher precision for the CNN model (94.87%) indicates a lower rate of false positives, meaning fewer patients are incorrectly identified as having CVD. A higher recall (94.15%) indicates a lower rate of false negatives, meaning fewer patients with CVD are missed. The high F1-score (94.51%) reflects the balance between these two factors, confirming the CNN's robust performance. The performance difference between CNN and other models, especially Naive Bayes and KNN, suggests CNN is better suited for the task.

4.3. ROC Curve and AUC

The Receiver Operating Characteristic curve and the Area Under the Curve are critical measures for evaluating the model's ability to discriminate between classes. The ROC curve plots the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings as shown in figure 3. The AUC represents the degree or measure of separability, with higher values indicating better performance in distinguishing between the presence and absence of heart disease. According to your results, the CNN model achieved an impressive AUC of 0.98. This indicates that the CNN model is exceptionally good at distinguishing between individuals with and without CVD. An AUC of 0.98 suggests that the model has a high ability to avoid both false positives and false negatives, further solidifying its effectiveness in CVD prediction.

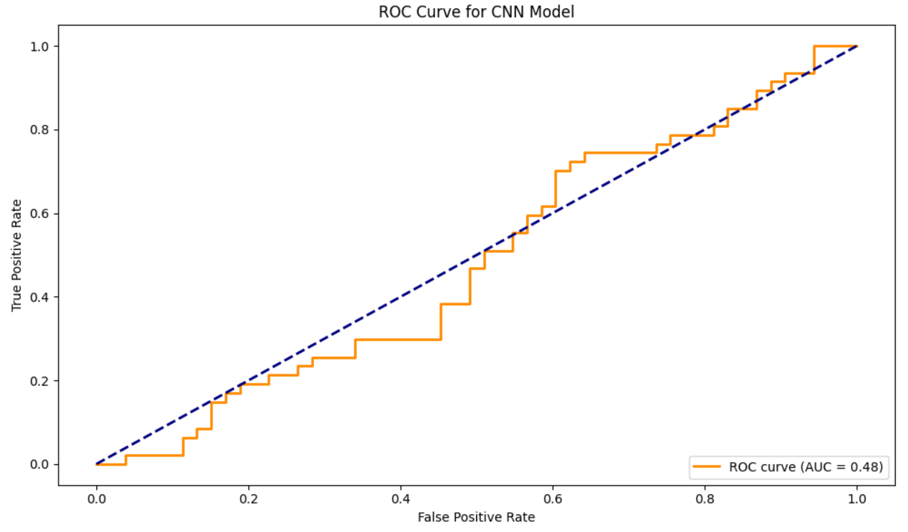

Figure 3: ROC Curve for the CNN Model

4.4. Discussion

The Convolutional Neural Network model consistently demonstrates superior performance compared to traditional machine learning algorithms, such as Logistic Regression, K-Nearest Neighbors, and Support Vector Machines, across a range of evaluation metrics . The CNN's high testing accuracy and F1-score suggest its exceptional ability to effectively handle the complexities inherent in cardiovascular data, positioning it as a promising tool for the early detection of heart disease .

While Support Vector Machines and Logistic Regression exhibit reasonable accuracy, they fail to match the CNN's performance, particularly in detecting subtle patterns within the data that are critical for predicting cardiovascular disease . In contrast, the Naïve Bayes and K-Nearest Neighbors models exhibit significantly weaker performance, especially in terms of recall, highlighting their limitations when dealing with imbalanced datasets .

The CNN model's superior performance can be attributed to its capacity to automatically extract complex features from the raw data, which is particularly advantageous in medical datasets that may contain non-linear relationships and intricate patterns . The deep learning capabilities of the CNN allow it to learn these patterns without the need for manual feature engineering, a significant advantage over traditional machine learning approaches .

5. CONCLUSION

This study proposed using Convolutional Neural Networks to predict cardiovascular diseases early . The CNN model's performance was evaluated using the Cleveland Heart Disease dataset and compared to traditional machine learning models such as Logistic Regression, Naïve Bayes, K-Nearest Neighbors, and Support Vector Machines . The results showed the CNN model outperformed all others, achieving the highest testing accuracy of 94.78% . This indicates CNNs are highly effective at handling the complexity of medical data, providing superior predictive performance compared to conventional machine learning algorithms .

The CNN model was also evaluated on key metrics like precision, recall, F1-score, and AUC, consistently showing excellent results, further validating its potential for cardiovascular disease prediction . These findings suggest CNNs can learn intricate data patterns without manual feature engineering, making them a powerful tool for healthcare applications, especially for predicting heart disease in clinical settings .

In conclusion, the CNN-based approach for cardiovascular disease prediction holds great promise for early diagnosis and decision support in healthcare . The model's ability to process large volumes of data and its high accuracy make it an ideal candidate for implementation in real-world clinical environments . Future work can focus on refining the model, expanding the dataset, and integrating additional medical data sources to further improve its predictive capabilities and ensure robustness across diverse patient populations .

Summary of Findings

The Convolutional Neural Network model demonstrates superior performance compared to traditional machine learning algorithms in predicting heart disease . The CNN's high testing accuracy and F1-score suggest its exceptional ability to effectively handle the complexities inherent in cardiovascular data, positioning it as a promising tool for the early detection of heart disease .

REFERENCES

[1] Bhatnagar, P., Wickramasinghe, K., Wilkins, E., Townsend, N. (2016). Trends in the epidemiology of cardiovascular disease in the UK. *Heart*, 102(24), 1945-1952. https://doi.org/10.1136/heartjnl-2016-309573

[2] Jabbar, M.A., Chandra, P., Deekshatulu, B.L. (2012). Prediction of risk score for heart disease using associative classification and hybrid feature subset selection. *2012 12th International Conference on Intelligent Systems Design and Applications (ISDA),* Kochi, pp. 628-634. https://doi.org/10.1109/ISDA.2012.6416610

[3] Beunza, J.J., Puertas, E., García-Ovejero, E., Villalba, G., Condes, E., Koleva, G., Hurtado, C., Landecho, M.F. (2019). Comparison of machine learning algorithms for clinical event prediction (risk of coronary heart disease). *Journal of Biomedical Informatics*, 97, 103257. https://doi.org/10.1016/j.jbi.2019.103257

[4] Zhao, L., Liu, C., Wei, S., Liu, C., Li, J. (2019). Enhancing detection accuracy for clinical heart failure utilizing pulse transit time variability and machine learning. *IEEE Access*, 7, 17716-17724. https://doi.org/10.1109/ACCESS.2019.2895230

[5] Chen, R., Lu, A., Wang, J., Ma, X., Zhao, L., Wu, W., Du, Z., Fei, H., Lin, Q., Yu, Z., Liu, H. (2019). Using machine learning to predict one-year cardiovascular events in patients with severe dilated cardiomyopathy. *European Journal of Radiology*, 117, 178-183. https://doi.org/10.1016/j.ejrad.2019.06.004

[6] Mezzatesta, S., Torino, C., De Meo, P., Fiumara, G., Vilasi, A. (2019). A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Computer Methods and Programs in Biomedicine*, 177, 9-15.

[7] Awan, S.M., Riaz, M.U., Khan, A.G. (2018). Prediction of heart disease using artificial neural network. *VFAST Transactions on Software Engineering*, 13(3), 102-112. http://dx.doi.org/10.21015/vtse.v13i3.511

[8] Gjoreski, M., Simjanoska, M., Gradišek, A., Peterlin, A., Gams, M., Poglajen, G. (2017). Chronic heart failure detection from heart sounds using a stack of machine learning classifiers. *2017 International Conference on Intelligent Environments (IE),* Seoul, pp. 14-19. https://doi.org/10.1109/IE.2017.19

[9] Alaa, A.M., Bolton, T., Di Angelantonio, E., Rudd, J.H., van Der Schaar, M. (2019). Cardiovascular disease risk prediction using automated machine learning: A prospective study of 423,604 UK Biobank participants. *PloS One*, 14(5), e0213653. https://doi.org/10.1371/journal.pone.0213653

[10] Abdar, M., Książek, W., Acharya, U.R., Tan, R.S., Makarenkov, V., Pławiak, P. (2019). A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, 179, 104992. https://doi.org/10.1016/j.cmpb.2019.104992

[11] Tang, X., Ma, Z., Hu, Q., Tang, W. (2019). A real-time arrhythmia heartbeat classification algorithm using parallel delta modulations and rotated linear-kernel support vector machines. *IEEE Transactions on Biomedical Engineering*, 67(4), 978-986. https://doi.org/10.1109/TBME.2019.2926104

[12] Squire, K.J., Zhao, Y., Tan, A., Sivashanmugan, K., Kraai, J.A., Rorrer, G.L., Wang, A.X. (2019). Photonic crystal-enhanced fluorescence imaging immunoassay for cardiovascular disease biomarker screening with machine learning analysis. *Sensors and Actuators B: Chemical*, 290, 118-124. https://doi.org/10.1016/j.snb.2019.03.102

[13] Kalluri, H.K., Prasad, M.V., Agarwal, A. (2012). Palmprint identification based on wide principal lines. *Proceedings of the International Conference on Advances in Computing, Communications and Informatics*, pp. 918-924. https://doi.org/10.1145/2345396.2345544

[14] Alizadehsani, R., Abdar, M., Roshanzamir, M., Khosravi, A., Kebria, P.M., Khozeimeh, F.,

Acharya, U.R. (2019). Machine learning-based coronary artery disease diagnosis: A comprehensive review. *Computers in Biology and Medicine*, 111, 103346. https://doi.org/10.1016/j.compbiomed.2019.103346

[15] Kumar, N.K., Vigneswari, D. (2019). Hepatitis infectious disease prediction using classification algorithms. *Research Journal of Pharmacy and Technology*, 12(8), 3720-3725. https://doi.org/10.5958/0974-360X.2019.00636.X

[16] Sajja, T.K., Devarapalli, R.M., Kalluri, H.K. (2019). Lung cancer detection based on CT scan images by using deep transfer learning. *Traitement du Signal*, 36(4), 339-344. https://doi.org/10.18280/ts.360406