Journal of Administrative and Economic Sciences • Vol 18, Issue 1 (2025) • p 477-491

DeepHeart: A Deep Learning Model for Predicting Heart Diseases Using Convolutional Neural Network

Ibrahim Almubark

Department of Information Technology, College of Computer, Qassim University, Buraydah • Saudi Arabia

Imbark@qu.edu.sa

Abstract:

Objectives: Heart disease is a significant and widespread cause of illness and death around the world. Traditionally, heart disease risk assessment relies on clinical variables. The integration of deep learning techniques in this domain is a challenging, yet promising, solution. **Methods:** In this study, we employed a Deep Convolutional Neural Network (DCNN) to predict the risk of heart disease using clinical cardiology data. The dataset used in the analysis was obtained from the publicly available UCI Heart Diseases Repository consisting of 920 patients' records comprising 14 attributes. The model was trained and fine-tuned using a grid search approach to optimize the best hyperparameters. **Findings:** The performance of the model was rigorously assessed through cross-validation techniques. We found that the proposed technique demonstrated remarkable performance, achieving an impressive testing accuracy of 83%. Precision, recall, and F1-score were equally notable at 84%, 83%, and 83%, respectively; showcasing the model's well-balanced classification capabilities. Our research produced promising results while highlighting the potential of the proposed DCNN model as a robust tool for heart disease prediction.

Keywords: deep convolutional neural network; deep learning; disease prediction and modeling; computer vision; heart disease

1. Introduction

Heart disease is a critical health problem all over the world, contributing significantly to morbidity and mortality. According to World Health Organization (WHO) statistics, cardiovascular diseases are the leading cause of death worldwide [1]. The impact of heart disease extends beyond mortality, affecting the quality of life of individuals and their families and posing a substantial economic burden on healthcare systems [2]. Traditional diagnostic methods for heart disease, including clinical assessments, medical imaging, and blood tests, play a crucial role in identifying cardiovascular diseases. Clinical assessments involve checking blood pressure and pulse rate to detect the heart failure tendency. Medical imaging also plays an essential role in identifying various heart diseases. Health-care professionals can visualize the internal functionalities of the body leading to timely detection and treatment of heart diseases [3]. Moreover, using medical images, doctors can examine the inner body structure without invasive procedures like biopsies or surgeries. Doctors also refer multiple tests such as Electrocardiograms (ECG), echocardiograms, Angiography, and blood tests to diagnose heart diseases. They use the results of these tests frequently by employing rule-based systems or risk calculators, to make precise diagnoses [4].

However, these methods have inherent limitations such as performance analysis, high costs, and limited accessibility, particularly in resource-constrained areas [5]. Hence, the need for more efficient and accessible predictive and diagnostic tools is evident. Machine learning (ML) and Deep Learning (DL) techniques have revolutionized healthcare by enhancing accuracy and efficiency in disease diagnosis [6]. These techniques have demonstrated the potential to analyze large datasets, identify patterns, and make predictions in various healthcare domains such as neurology, nephrology, and cardiology [7]. This has enabled healthcare professionals to increase the speed and accuracy of the disease diagnosis and improve decisions related to treatment schedules [8]. This has led to improved patient outcomes, personalized treatment plans, and more timely interventions at reduced costs to the individual patient [9].

Despite the advancements in ML and DL-based applications for healthcare, existing techniques for heart disease diagnosis may face challenges. These challenges include the interpretability of the results, potential biases in training data, and the need for more robust and generalizable approaches, especially when dealing with diverse patient populations.

The primary purpose of this study is to leverage DL techniques in the prediction of heart diseases. In addition, we aim to enhance the performance and efficiency of heart disease prediction while considering a diverse set of patients and their unique characteristics. We propose a novel approach to heart disease prediction using a Deep Convolutional Neural Network (DCNN). By utilizing the publicly available UCI Heart Diseases Repository datasets comprising 920 patients' records with 14 attributes, we have employed and optimized DCNN. The innovation lies in our approach to discriminating between healthy and cardiac disease patients. To evaluate the performance of the model, we used several matrices including accuracy, precision, recall (sensitivity), F1-score, and Area Under the Curve AUC [10], [11]. The experimental results show the effectiveness of the proposed method in a real-world

environment. The following points can further highlight our main contributions to the research:

- **Specialized DCNN for Heart Disease Prediction:** Our study presents a specialized DCNN specifically designed for heart disease prediction, demonstrating excellent performance in classification tasks.
- **Advanced Pre-processing Techniques:** This study includes advanced pre-processing techniques such as handling missing data, normalizing values, and feature selection, as well as a DCNN with hyperparameter tuning using a grid search method.
- **Resilience and Generalizability:** Our study aims to ensure resilience and generalizability to different patient populations. The assessment includes an extensive set of performance measures such as accuracy, precision, recall, F1-score, and AUC providing comprehensive evaluations instead of relying on a single measure to ensure model reliability.
- **Timely Interventions and Personalized Treatments:** This facilitates timely interventions and personalized treatment plans by improving the accuracy of diagnosing heart diseases, thereby bridging a significant gap in the current healthcare systems.

2. Related Work

Various studies have used ML and DL to predict diseases, with findings varying greatly depending on the researchers' techniques.

In a study using the Cleveland heart disease dataset [12], the results showed that logistic regression was the most effective algorithm, with an accuracy of 82.89%, followed by Support Vector Machine (SVM) at 81.57%, and both decision tree and Naïve Bayes at 80.43%. This study demonstrated how ML can be used to increase the precision of heart disease diagnosis and early detection.

In another study by Bharti, Rohit, *et al*. [13], different ML and DL models were used, including neural networks, SVM, decision tree, random forest, and CART algorithms for heart disease prediction. Feature engineering and feature selection were employed to improve both classification and prediction. Random forest achieved 80.3% accuracy, while decision tree had an accuracy of 82.3% based on features such as blood sugar level, age of the patient, and blood pressure.

Voon Khai Tick *et al.* [14], investigated the classification of heart disease using ML approaches, with a particular focus on Artificial Neural Networks (ANN). They used back-propagation training to train a multi-layered perceptron to classify data on heart illness, including cases with and without heart attacks. The study investigated several learning rates and neuron configurations using sigmoid activation functions across 1000 epochs. The highest result was obtained with a learning rate of 0.25 and 25 neurons, producing an accuracy of 80.66%. This study showed how well ANN can classify cases of heart disease and identify data trends.

Jyotismita Talukdar and Thipendra P. Singh [15], investigated the use of ANN to predict cardiovascular disease, a primary cause of death in India. The importance of leveraging large volumes of healthcare data to cut costs and enhance diagnosis accuracy was underlined. The paper examined ML approaches, identified effective techniques for predicting and classifying cardiac diseases, and presented a neural network model for identifying risk factors by examining correlations between them. Using back-propagation, the model attained an accuracy of 81%. This technique aims to improve decision-support systems and reduce medical errors.

Kim and Kang [16], investigated neural network (NN)-based coronary heart disease (CHD) risk prediction through feature correlation analysis (NN-FCA). The study addressed the limits of NN's "black-box" nature by implementing NN-FCA, a two-stage approach that includes feature selection and correlation analysis. The study evaluated 4146 individuals from a Korean dataset and found that NN-FCA had a high accuracy of 82.51% in CHD prediction, outperforming the Framingham risk score (FRS). The area under the receiver operating characteristic (ROC) curve for NN-FCA (0.749 ± 0.010) outperformed that of FRS (0.393 ± 0.010), indicating better predictive potential. These findings demonstrated the effectiveness of NN-FCA in improving CHD risk prediction accuracy compared to older approaches.

Fazl-e-Rabbi *et al.* [17], utilized several classifiers to predict cardiac disease using the Cleveland dataset from the UCI repository, which included 270 records with 76 features. This study used only 13 of the dataset's features. Three distinct classifiers were employed to predict cardiac diseases: SVM, ANN, and k-nearest neighbor. The classification accuracy using SVM machines was 85.18%. The accuracy of the k-nearest neighbor continued to rise as the number of k increased until $k = 10$, at which point it achieved an accuracy of 80.74%. The accuracy of the ANN was 73.33%.

M. Bhagawati *et al.* [18], reviewed DL techniques specifically solo DL and hybrid DL models for cardiovascular disease risk classification. They examined 286 studies while discussing DL architecture, their implementation, and bias reduction mechanisms. They explored electrocardiogram-based solutions and bias assessment tools such as PROBAST and ROBINS-I. The results revealed that ensemble-based techniques could outperform solo and hybrid DL methods due to their accuracy, speed, and reduced bias.

L. Luo *et al.* [19], designed a deep-learning ensemble method for automatic identification of heart-disease risk factors from health records. This method removed the need for external domain knowledge by using Bidirectional Encoder Representations from Transformers (BERT) to extract meaningful features from clinical notes. These features were then analyzed by conditional random fields to detect risk-factor identifiers. Experimental results presented that the proposed approach attained state-of-the-art performance in identifying risk factors.

L. Mhamdi *et al.* [20], designed an algorithmic model for the prediction of cardiovascular diseases from electrocardiogram tracings. The study utilized MobileNetV2 and VGG16 algorithms which improved medical diagnosis at reduced cost. It implemented several optimization techniques and achieved a validation accuracy of 95% for both MobileNetV2 and VGG16 algorithms. However, this

accuracy decreased to 94% for MobileNetV2 and 90% for VGG16 when these algorithms were implemented using Raspberry Pi.

I. T. Joseph S *et al.* [17], discussed the frequent attack of heart diseases in the modern world due to the lack of healthcare awareness. They emphasized the need for effective techniques to diagnose heart disease symptoms timely leading to early treatment. They highlighted that ML and DL techniques had the potential to correct predictions of heart diseases by employing large datasets. The study explored various ML and DL models that could identify heart diseases using publicly available benchmark datasets.

A summary of the related work section is given in Table 1.

Table 1. Related work summary.

Study	Technique used	Key Findings	Accuracy
Aburayya, Rand Abedelellah, *et al.* [12]	Logistic regression, SVM, decision tree, and Naïve Bayes	Logistic regression outperformed other models	Logistic regression: 82.89%
Bharti, Rohit, *et al.* [13]	Neural networks, SVM, decision tree, random forest, and CART	Decision tree outperformed other models	Decision tree: 82.3%,
Voon Khai Tick *et al.* [14]	Multi-layer perceptron (MLP)	Best accuracy achieved with a learning rate of 0.25 and 25 neurons	MLP: 80.66%
Talukdar, Jyotismita , and Thipendra P. Singh [15]	ANN	Focused on cardiovascular disease by highlighting the importance of large healthcare datasets	ANN: 81%
Kim, Jae Kwon, and Sanggil Kang [16]	Neural networks (NN-FCA)	NN-FCA outperformed traditional risk scores in coronary heart disease prediction	NN-FCA: 82.51%
Rabbi, Md Fazle, *et al.* [17]	SVM, ANN, and k-nearest neighbor	SVM showed the highest accuracy in predicting cardiac diseases	SVM: 85.18%
M. Bhagawati *et al.* [18]	DL, hybrid deep learning, and ensemble-based methods	Ensemble-based methods outperformed solo and hybrid DL approaches	Not specified
L. Luo *et al.* [19]	BERT	Achieved state-of-the-art performance	Not specified

		in identifying heart disease risk	
L. Mhamdi *et al.* [20]	MobileNetV2 and VGG16	High validation accuracy for both MobileNetV2 and VGG16	MobileNetV2, VGG16: 95%
I. T. Joseph S *et al.* [21]	ML and DL models	Emphasized the potential of ML/DL for early heart disease diagnosis	Not specified

3. Materials and Methods

In this study, we have employed a Convolutional Neural Network (CNN) for the prediction of heart disease. The process started with the collection of patients' data. The data is gathered and enters the pre-processing phase. Several pre-possessing techniques were applied to the dataset for its improvement – including handling the missing values, normalization, and selecting the best features. The cleaned dataset was then divided into two subsets: (1) the training dataset and (2) the testing dataset. The training set was utilized to train the model and later the test set was used to evaluate the performance of the model. The proposed model categorizes the individual into either healthy or heart disease class. A step-by-step approach employed for deep learning-based heart disease prediction is shown in Figure 1.

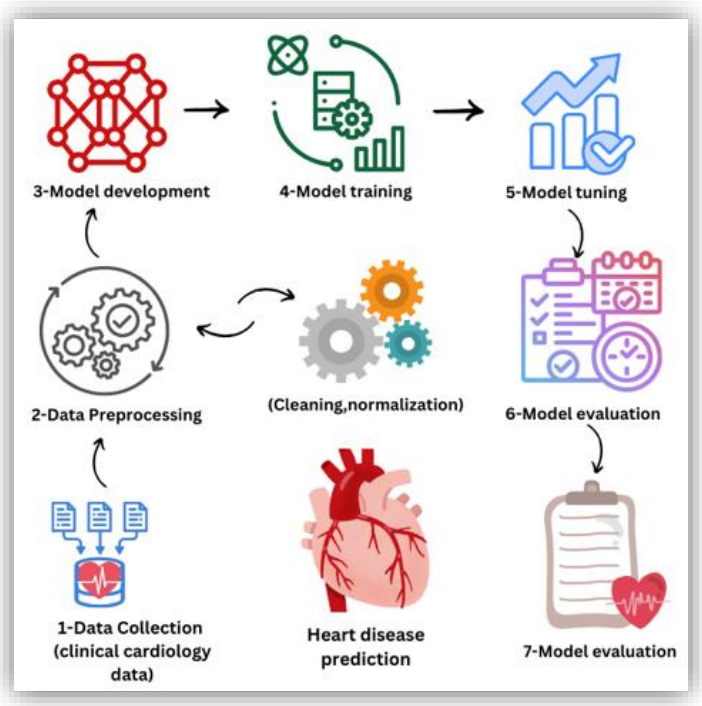

Figure 1. Deep learning-based heart disease prediction.

3.1 Dataset

The dataset of heart disease patients used for the proposed methodology is gathered from the University of California (UCI, Irvine C.A) repository, which is publicly

available on the Kaggle website[1]. The dataset has been widely used for classification tasks in the field of health and medicine. There are several other datasets available for the same task, such as the Heart Disease dataset [22] and the Cleveland Clinic Heart Disease dataset [23]. The Cleveland dataset consists of records from patients who belong to Cleveland, while the Heart Disease dataset is more diverse, comprising records collected from various locations including Hungary, Switzerland, Long Beach V, and Cleveland itself. The Heart Disease dataset is a pre-processed version of the UCI dataset, which can help skip the pre-processing steps. However, to ensure the correct processing of the DCNN model, we will utilize the original UCI dataset and apply preprocessing techniques on it. The dataset contains a total of 920 records of the patients with 16 different attributes of which 14 were used in this study. The 'Heart disease' attribute in the dataset which predicts the presence of heart disease in a patient, is represented by an integer value that lies in the domain [0,4] (0 representing the healthy case and 1 to 4 representing the different stages of the heart disease found in the patient record). Table 2 shows the different attributes of the dataset along with their description.

Table 2. Dataset attributes and variable description.

No.	Attribute	Representation	Description
1	ID	id	Unique row number
2	Age	age	Age in years
3	Location	dataset	City name
4	Gender	sex	Male and female
5	Chest pain	cp	Four types of chest pain
6	Cholesterol level	chol	Measure of cholesterol in mg/dl
7	Resting blood pressure	trestbps	Blood pressure when the body is in a state of rest
8	Fasting blood sugar	fbs	Blood sugar level while fasting
9	MaxHR	thalach	Maximal heart rate
10	Resting ECG	restecg	Resting electrocardiograph
11	Exercise-induced angina	exang	Exercise-induced angina
12	Old peak	oldpeak	ST depression brought by exercise comparative to rest
13	Slope	slope	Slope of exercise peak
14	Vessels	ca	No. of major vessels
15	Thalassemia	thal	Normal, fixed, and reversible defects
16	Heart disease	num	Predicted attribute

[1] https://www.kaggle.com/datasets/redwankarimsony/heart-disease-data

3.2 Data Pre-processing

In this study, we focused on a binary classification identifying whether a patient had heart disease (combining stages 1-4) or not (stage 0). To simplify the analysis, we merged all four heart disease stages into a single "heart disease" class 1 (509 instances), while stage 0 remained the "healthy" class (411 instances).

Certain attributes with a high percentage of missing values, namely ID, location, slope, vessels, and the presence of thalassemia (yes/no), were excluded from the dataset. The decision to omit certain attributes was guided by various considerations. The attribute labeled ‘Vessels’ was excluded due to its high proportion of null values (exceeding 60%). Additionally, the attributes ‘Slope’ and ‘Thalassemia’ were excluded because analysis showed their correlation with the target variable was low (between 30% - 35%). Finally, attributes ‘ID’ and ‘Location’ were excluded as their relevance to model training purposes was limited. After deleting these attributes, the dataset was left with 11 attributes.

To handle the Not a Number (*NaN*) values, the mean imputation technique was used to impute the null values in the attributes (resting blood pressure (missing 59), MaxHR (missing 55), and cholesterol level (missing 30)), while the mode imputation technique was employed to handle the attributes (exercise-induced angina (missing 55)). After the imputation technique, the dataset had only two *NaN* values left in the ‘Resting ECG’ column and it is the most common approach to drop the rows when they are very small in number as compared with the total sample size, so following this approach, we dropped the *NaN* rows from the dataset. At the end of this step, the dataset was left with 0 *NaN* values.

During the data encoding phase, categorical columns such as gender and fasting blood sugar were transformed. The gender column had two values (Male and Female), with Male mapped to value 1 and Female to 0. Similarly, fasting blood sugar had two boolean values (True or False), with True mapped to value 1 and False to 0. The chest pain and resting ECG columns were transformed using one-hot encoding which increased the number of total features back to 14. To ensure consistent scaling, MinMaxScaler was applied to normalize the data in columns for age, resting blood pressure, cholesterol level, and MaxHR across the entire dataset.

After completing the data pre-processing steps, we had a total of 918 instances. Among these, 411 instances belonged to class 0 (indicating no heart disease), and 509 instances belonged to class 1 (indicating the presence of heart disease). The dataset was then divided into two subsets, the first was the training set which included 80% of the dataset and the remaining 20% was considered to be the test dataset.

3.3 Deep Learning Model Architecture

In this work, we have employed a Deep Convolutional Neural Network (DCNN) architecture. The architecture begins with an input shape of (14, 1) which undergoes convolutional operation with 256 filters. Each one with a kernel size of 3 × 3, a stride of one, a Rectified Linear Unit (ReLU) activation function, and SAME padding. Additionally, *L2* regularization is applied to the convolutional layer to avoid overfitting.

Batch normalization is then performed to normalize the activations of the convolutional layer. This is followed by max pooling with a pool size of two to downsample the feature maps.

The subsequent layer comprises a convolutional operation with 512 filters, each incrementally larger than the previous layer, followed by batch normalization and dropout regularization with a rate of 20% to prevent overfitting. The output of this convolutional layer is flattened using a flattened layer, converting the multi-dimensional feature maps into a one-dimensional tensor.

The flattened output is then fed into a series of fully connected layers, consisting of units (128, 256, 512, 1024, 2048, 1024, 512), respectively. Each fully connected layer, apart from the last one, is accompanied by a ReLU activation function, batch normalization, and dropout regularization with a rate of 20%. The final layer is a dense layer with a single unit and sigmoid activation function, serving as the output layer to produce the model's predictions. The formula for a sigmoid function is described here.

$$S(x) = \frac{1}{1 + e^{-x}}$$

Our DCNN employed the sigmoid function, a mathematical tool shaped like an "*S*" curve [24]. This function excels at transforming any numerical value into a probability score between 0 and 1 [25]. In our DCNN, it resided in the final layer, converting the raw output into a clear probability indicating the presence or absence of heart disease. Figure 2 shows the architecture of the proposed DCNN.

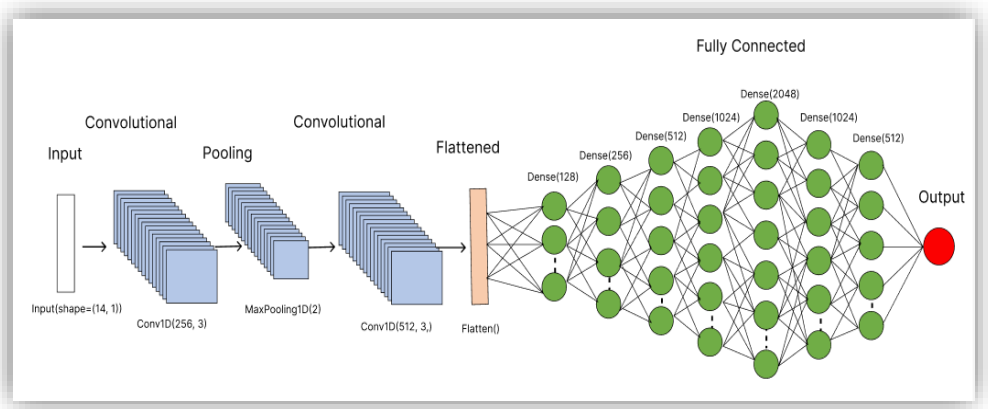

Figure 2. The model architecture used in the proposed DCNN system.

During training, the model underwent 100 epochs using a batch size of eight and was optimized using the Adam optimizer. Binary cross-entropy was employed as the loss function, while training progress was tracked using the accuracy metric. Hyperparameters, crucial for model generalization, were carefully tuned throughout training. Both *L1* and *L2* regularization were applied to each layer to mitigate overfitting risks. Dropout layers randomly removed neurons from dense layers, providing an additional defense against overfitting. The training included an early

stopping mechanism, which stopped the process if no improvement in accuracy was observed for 15 consecutive epochs.

Additionally, the model checkpoints were saved, preserving weights that outperformed previous validation accuracies for potential use post-training. The inclusion of a validation set, unseen by the model during training, aided in evaluating performance and effectively identifying potential overfitting. These measures collectively enhanced the robustness of the model and prevented performance degradation on test data.

3.4 Evaluating the Performance of the Model

The performance of the proposed model was evaluated by using several performance metrics such as accuracy, precision, recall, F1 score, and AUC. In these measures, True Positive (TP) shows the correctly classified heart disease patients, while True Negative (TN) represents the correctly classified healthy patients. On the other hand, False Positive (FP) and False Negative (FN) show the incorrect classification of healthy and heart disease patients respectively.

3.4.1. Accuracy

One way to measure how frequently a DL model correctly classifies a data point is to evaluate the accuracy of the algorithm, which is the ratio between the number of correct predictions over the total prediction. The accuracy can be calculated using Equation 1:

$$Accuracy = \frac{TP + TN}{(TP + FP + TN + FN)} \quad (1)$$

3.4.2. Precision

Precision (also known as positive predictive value) is the measure of the number of correct predictions to the total number of predicted positives. The precision can be calculated using Equation 2:

$$Precision = \frac{T\,P}{TP + FP} \quad (2)$$

3.4.3. Recall

Recall (also known as sensitivity) represents the ratio of correct class predictions to the total number of actual positives in the dataset [26]. It is a measure to determine the completeness of the classifier. The recall can be calculated using Equation 3:

$$Recall = \frac{TP}{TP + F\,N} \quad (3)$$

3.4.4. F1-Score

It can sometimes be difficult to decide whether high precision or low recall is better, or vice versa when comparing different models. The combination of precision and recall is called the F1 score [27]. It can be measured using equation 4:

$$F1\ Score = \frac{2 \times (Precision \times Recall)}{(Precision + Recall)} \quad (4)$$

3.4.5 AUC

It represents how well the model is trained to classify the data using various threshold levels. Its value is between 0 and 1. The closer the results are to 1, the better the model performs.

4. Results and Discussions

The proposed DCNN model showed excellent performance in predicting and classifying cases of heart disease as shown in Table 3. The model achieved an impressive accuracy of 83%. Precision, recall, and F1 scores were equally high at 84%, 83%, and 83%, respectively. Furthermore, evaluating the model's performance highlighted its robustness, demonstrating its reliability. However, it's important to note some limitations of the proposed model, such as occasional inaccurate predictions, suggesting the need for further improvement in its predictive capabilities. In summary, the DCNN model shows promising results in accurately predicting heart disease, but there is room for enhancement through continued research.

Table 3. Performance matrix for the DCNN model.

Performance matrix	Performance %
Accuracy	83
Precision	84
Recall	83
F1-Score	83

The confusion matrix provided in Figure 3 serves to assess the performance of the proposed DCNN model. The diagonal elements represent the True Negative (55) and True Positive (56) predictions made by the model. However, the model misclassified 15 instances as false positives and 7 instances as false negatives.

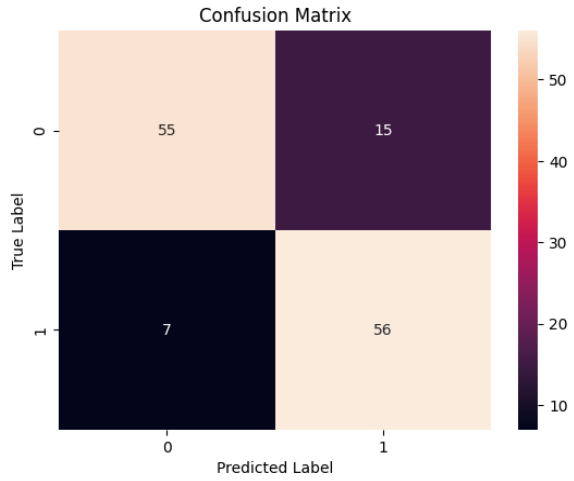

Figure 3. The confusion matrix of the proposed model.

The performance of the classification algorithms is intrinsically linked with the area under the curve (AUC), i.e., the larger the value of the AUC is, the better the performance of the classification algorithm. The proposed model's strength and robustness were evaluated using the ROC curve. The receiver operator characteristics curve of the proposed model is shown in Figure 4.

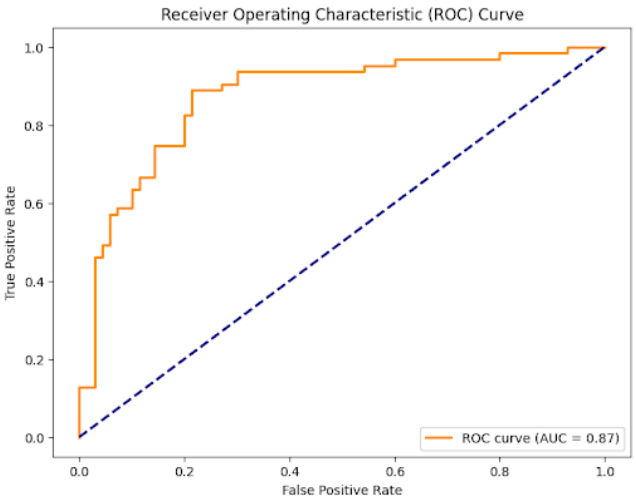

Figure 4. ROC curve of the proposed on the proposed model.

Our study unveils a groundbreaking advancement in heart disease prediction through the implementation of the DCNN model. The model showcased exceptional accuracy, robustness, and potential for clinical application. With an accuracy rate of 83%, our DCNN model not only surpassed traditional methods but also demonstrated high precision and recall rates of 84% and 83%, respectively, ensuring reliable positive and negative identifications. Moreover, its balanced F1 score solidifies its overall effectiveness. Moreover, the confusion matrix unveils the in-depth workings of the model, highlighting its adaptiveness in classifying true positives and true negatives while managing potential error cases. While our model shares similarities with previous DL approaches in predicting early heart disease, its distinguishing feature lies in its superior accuracy and efficiency, outperforming prior models by significant margins. However, we acknowledge that reliance on extensive data and computational resources for optimal performance, within the model, poses a challenge – particularly in resource-limited settings. Despite these limitations, the strengths of our DCNN model far outweigh its drawbacks. Looking ahead, our research paves the way for future improvements to leverage larger datasets and diverse DCNN architectures to enhance predictive capabilities further. Exploring variations in architecture, optimization techniques, loss functions, and fine-tuning parameters holds promise for unraveling new insights and optimizing model performance.

5. Conclusions

Heart disease prediction at an early stage increases treatment efficacy and, thus, reduces the incidence of morbidity and mortality. In healthcare, increasingly machine learning

and deep learning models are being used to predict the risk and presence of disease accurately. Our study utilized patient data and the DCNN model to demonstrate how the use of an efficient algorithm can help physicians detect the possible presence of heart disease before it manifests. In the future, the proposed model can be used to assist in timely diagnosis and, therefore, support treatment decisions. In the future, more research should be conducted to refine the DCNN model, including larger datasets, various patient populations, and advanced optimization approaches.

Acknowledgments: The author would like to thank Qassim University and the Deanship of Scientific Research for their support and funding the publication.

References:

[1] World Health Organization, "Cardiovascular diseases." Accessed: Sep. 06, 2024. [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases

[2] F. Sun *et al.*, "Social Determinants, Cardiovascular Disease, and Health Care Cost: A Nationwide Study in the United States Using Machine Learning," *Journal of the American Heart Association*, vol. 12, no. 5, p. e027919, Mar. 2023, doi: 10.1161/JAHA.122.027919.

[3] B. Barone Gibbs *et al.*, "Physical Activity as a Critical Component of First-Line Treatment for Elevated Blood Pressure or Cholesterol: Who, What, and How?: A Scientific Statement From the American Heart Association," *Hypertension*, vol. 78, no. 2, Aug. 2021, doi: 10.1161/HYP.0000000000000196.

[4] M. A. Khan and F. Algarni, "A Healthcare Monitoring System for the Diagnosis of Heart Disease in the IoMT Cloud Environment Using MSSO-ANFIS," *IEEE Access*, vol. 8, pp. 122259–122269, 2020, doi: 10.1109/ACCESS.2020.3006424.

[5] M. M. R. Khan Mamun and T. Elfouly, "Detection of Cardiovascular Disease from Clinical Parameters Using a One-Dimensional Convolutional Neural Network," *Bioengineering (Basel)*, vol. 10, no. 7, p. 796, Jul. 2023, doi: 10.3390/bioengineering10070796.

[6] F. M. Delpino, Â. K. Costa, S. R. Farias, A. D. P. Chiavegatto Filho, R. A. Arcêncio, and B. P. Nunes, "Machine learning for predicting chronic diseases: a systematic review," *Public Health*, vol. 205, pp. 14–25, Apr. 2022, doi: 10.1016/j.puhe.2022.01.007.

[7] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey," *Journal of Electrical Systems and Inf Technol*, vol. 10, no. 1, p. 40, Aug. 2023, doi: 10.1186/s43067-023-00108-y.

[8] M. Javaid, A. Haleem, R. Pratap Singh, R. Suman, and S. Rab, "Significance of machine learning in healthcare: Features, pillars and applications," *International Journal of Intelligent Networks*, vol. 3, pp. 58–73, Jan. 2022, doi: 10.1016/j.ijin.2022.05.002.

[9] A. S. Ahuja, "The impact of artificial intelligence in medicine on the future role of the physician," *PeerJ*, vol. 7, p. e7702, Oct. 2019, doi: 10.7717/peerj.7702.

[10] M. Ali *et al.*, "Software Defect Prediction Using an Intelligent Ensemble-Based Model," *IEEE Access*, vol. 12, pp. 20376–20395, 2024, doi: 10.1109/ACCESS.2024.3358201.

[11] M. Ali, T. Mazhar, A. Al-Rasheed, T. Shahzad, Y. Yasin Ghadi, and M. Amir Khan, "Enhancing software defect prediction: a framework with improved feature selection and ensemble machine learning," *PeerJ Computer Science*, vol. 10, p. e1860, Feb. 2024, doi: 10.7717/peerj-cs.1860.

[12] R. A. Aburayya *et al.*, "Automated Heart Diseases Detection Using Machine Learning Approach," in *2023 6th International Conference on Engineering Technology and its Applications (IICETA)*, Jul. 2023, pp. 108–114. doi: 10.1109/IICETA57613.2023.10351330.

[13] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," *Computational Intelligence and Neuroscience*, vol. 2021, no. 1, p. 8387680, 2021, doi: 10.1155/2021/8387680.

[14] V. K. Tick, N. Y. Meeng, N. F. Mohammad, N. H. Harun, H. Alquran, and M. F. M. Mohsin, "Classification of Heart Disease using Artificial Neural Network," *J. Phys.: Conf. Ser.*, vol. 1997, no. 1, p. 012022, Aug. 2021, doi: 10.1088/1742-6596/1997/1/012022.

[15] J. Talukdar and T. P. Singh, "Early prediction of cardiovascular disease using artificial neural network," *Paladyn, Journal of Behavioral Robotics*, vol. 14, no. 1, Jan. 2023, doi: 10.1515/pjbr-2022-0107.

[16] J. K. Kim and S. Kang, "Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis," *Journal of Healthcare Engineering*, vol. 2017, no. 1, p. 2780501, 2017, doi: 10.1155/2017/2780501.

[17] Md. F. Rabbi *et al.*, "Performance Evaluation of Data Mining Classification Techniques for Heart Disease Prediction," *Journal of Engineering Research*, Feb. 2018.

[18] M. Bhagawati *et al.*, "Cardiovascular disease/stroke risk stratification in deep learning framework: a review," *Cardiovasc Diagn Ther*, vol. 13, no. 3, pp. 557–598, Jun. 2023, doi: 10.21037/cdt-22-438.

[19] L. Luo, Y. Wang, and D. Y. Mo, "Identifying heart disease risk factors from electronic health records using an ensemble of deep learning method," *IISE Transactions on Healthcare Systems Engineering*, vol. 13, no. 3, pp. 237–247, Jul. 2023, doi: 10.1080/24725579.2023.2205665.

[20] L. Mhamdi, O. Dammak, F. Cottin, and I. B. Dhaou, "Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems," *Biomedicines*, vol. 10, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/biomedicines10082013.

[21] I. T. Joseph S, V. S, M. M, P. L. Jancy, and K. P. A, "ML & DL Methodologies in Heart Disease Prediction: Brief Analysis," in *2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2022, pp. 208–212. doi: 10.1109/ICACCS54159.2022.9785241.

[22] "Heart Disease Dataset." Kaggle. Accessed: Sep. 06, 2024. [Online]. Available: https://www.kaggle.com/datasets/johnsmith88/heart-disease-dataset

[23] "Cleveland Clinic Heart Disease Dataset." Accessed: Sep. 06, 2024. [Online]. Available: https://www.kaggle.com/datasets/aavigan/cleveland-clinic-heart-disease-dataset

[24] A. I. Rodríguez and X. D. Buitrago, "How to choose an activation function for deep learning," *Tekhnê*, vol. 19, no. 1, pp. 23–32, May 2022.

[25] A. Paul, R. Bandyopadhyay, J. H. Yoon, Z. W. Geem, and R. Sarkar, "SinLU: Sinu-Sigmoidal Linear Unit," *Mathematics*, vol. 10, no. 3, p. 337, Jan. 2022, doi: 10.3390/math10030337.

[26] M. Ali *et al.*, "Analysis of Feature Selection Methods in Software Defect Prediction Models," *IEEE Access*, vol. 11, pp. 145954–145974, 2023, doi: 10.1109/ACCESS.2023.3343249.

[27] M. Ali, M. MSohaib Azam, and T. Shahzad, "Random Search-Based Parameter Optimization on Binary Classifiers for Software Defect Prediction," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)*, vol. 10, no. 2, pp. 476–488, 2024.