

# Heart Disease Prediction Using Convolutional Neural Network

Bhanumathi M<sup>a,1</sup>, Gautham Kumar RS<sup>b</sup>, Karthika Hema Manasa M<sup>b</sup> and Aravindh R<sup>b</sup>

<sup>a</sup>Assistant professor, Department of Computer Science Engineering, Easwari Engineering College, Chennai, India

<sup>b</sup>Undergraduate students, Department of Computer Science Engineering, Easwari Engineering College Chennai, India

**Abstract.** The group of diseases related to heart and blood vessels are commonly known as Cardio vascular diseases. Considering today's lifestyle, they are the major cause of deaths globally. So early identification of the diseases can give us a fair chance of effective treatment and proper precautions for a healthy life. The proposed novel model is created using convolutional neural networks in combination with random forest algorithm and there will be 12 factors that help to predict the heart disease and chest X-rays will also be used as one of the parameters.

**Keywords.** Cardiovascular diseases, convolutional neural networks, Random forest algorithm.

## 1. Introduction

With digitalization around the world, technology plays a massive role in our daily tasks. But today's technological advancements implemented in the healthcare sector are dominating small devices to huge equipment. The rapid introduction of technology in the medical domain can deliver an accurate diagnosis, promising cure and effective medications and treatments with an increased preciseness. According to WHO, Cardiovascular diseases are the major cause of death globally. In a recent survey, an approximate of 17.9 million premature people (under the age of 70) died due to the heart diseases. In the proposed novel method, we are focused on getting an accurate prediction of the heart disease of the user using Convolutional neural networks (which are known for its human-like processing) and the classification model. Heart diseases like Arrhythmias, Cardiomyopathy, Ischemic heart disease, inflammatory heart disease can be identified using their respective symptoms [1–7]. The input from the user will be preprocessed, tested and an appropriate report will be generated with the necessary steps that are needed to be taken by the patient. The application can be deployed in the hospitals for a quicker and an accurate diagnosis instead of the manual diagnosis. The report can be used as a reliable source for the doctor in the further diagnosis.

---

<sup>1</sup>Bhanumathi M, Department of CSE, Easwari Engineering College, Chennai, India.  
E-mail: bhanuism@gmail.com.

## **2. Literature Survey**

Senthil Kumar et al. [8] provided a prediction model of hybridization of random forest with a linear model (HRFLM) was developed. It was a computer aided decision support system which could be used for the prediction(yes or no). HRFLM is a multilayer perception of neural networks. A combination of artificial neural network and back propagation technique and recurrent fuzzy networks were used to enhance accuracy in diagnosis.

Ram Prakash et al. [9] provided a prediction model of Deep Neural Network,  $\chi^2$  - statistical model where used to overcome the under fitting and over fitting issues faced in the previous model using artificial neural networks. The dataset was divided into training and testing datasets. As neural networks is made up of layers with the activation function. The input layer is fed with the training features and hidden layer is fed with the actual process data with weighed connection. Hypothesis was generated using the data in the trained model. Hypothesis was formulated with forward propagation .

The authors of [10] done research to improve the accuracy of the existing models of machine learning techniques naïve bayes, Random forest algorithm and support vector machine by combining it with the feature selection approach. An optimum amount of dataset was chosen, Rapid miner was used for data cleaning step. It was used to find the missing values. The outlier detection was applied along with rapid miner operator and distance method.

Maruf et al. [11] proposed an idea for prediction of heart diseases based on external factors such as taking drugs, smoking, diabetes, hereditary, shortness of breath, gender, unusual stress regularity of exercise, blood pressure, dizziness, age, tightness in chest using six distinct classifiers of ML (decision tree, random forest, naive bayes, support vector, quadratic discriminant, logistic regression). The confusion matrixes of prediction results were obtained. Support vector machine showed the best results.

## **3. System Architecture**

The architecture of proposed model using convolutional neural network consists many components. The first component is where the input for every prediction takes place; we need an input dataset based on which the prediction takes place for the given input. The patient details are converted in the form of dataset and it is combined with the input dataset both acts an input for data acquisition which is the process of the systems to perform the tasks such as conversion of data, storage of data, transmission of data and processing of data. Once data acquisition is over, the output of this is given to the pre-processing component where the data is processed and this result is given to the classification module. The chest X-ray which is taken as one of the attribute is processed using convolutional neural network. The classification module is built using Random Forest algorithm. All the outputs for the processing are fed into the module. The prediction of heart disease is done. The generated report can be a reliable document for the doctor for the further diagnosis.

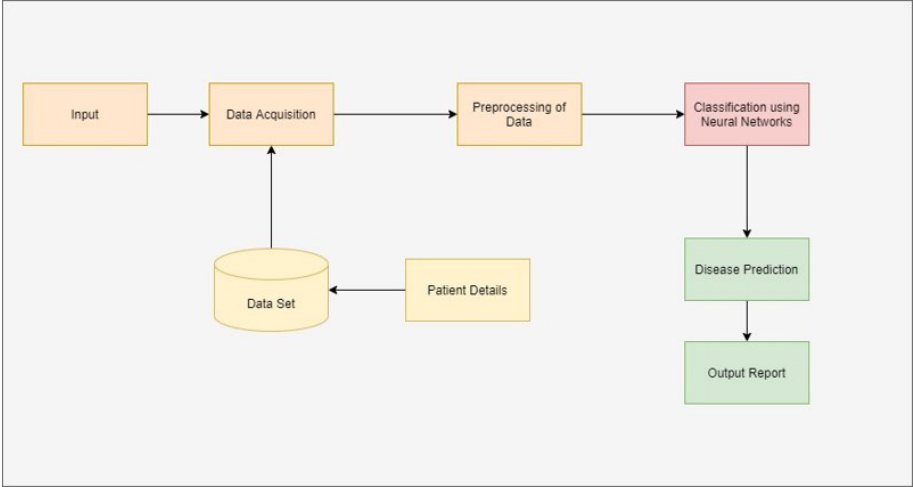


Figure 1. Heart Disease Prediction System

4. Implementation

4.1. GUI & Preprocessing Module

In this interface, the user or patient interacts with the application .This User interface is built in a simple way to make it user-friendly. This is the module in which the user enters the input parameters such as Age, Triceps skin fold thickness, Plasma glucose concentration, Insulin, Diabetes pedigree function(DPI), Diastolic blood pressure, Number of times pregnant, Body mass index, FBS, Cp, Old Peak, Thal and X-ray image. The dataset is obtained from the UCI repository 90% data is utilized in training the machine and 10% is utilized used to test the machine.

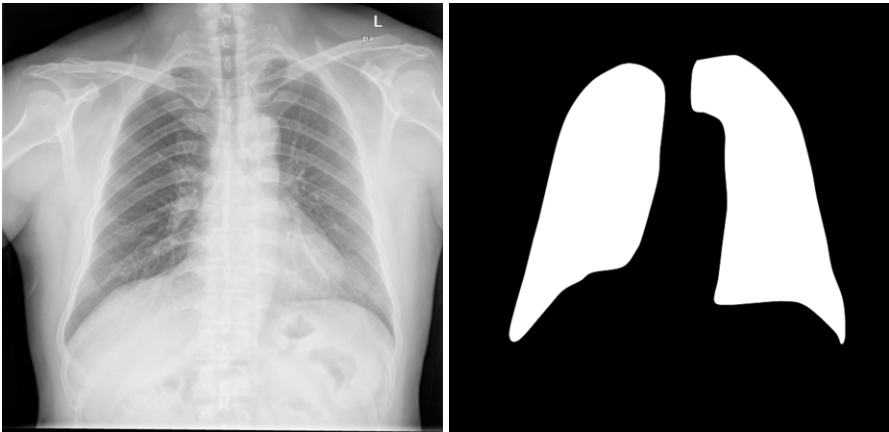
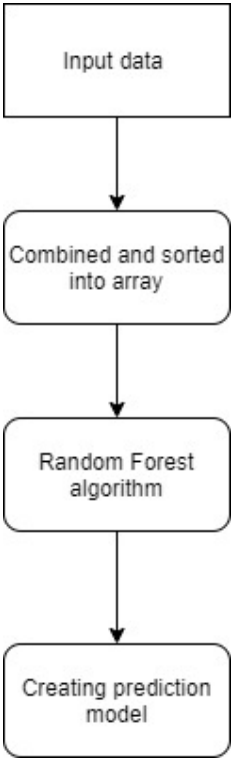


Figure 2. X-ray Image Before and After Processing

Data preprocessing is done where data is cleaned and missing values are filled in by the mean value to avoid mishandles in the prediction. A 2D Convolutional neural network

is created using `keras.layers.convolutional`. The chest X-ray obtained from the user is also processed by convolutional neural network as it is best suited for the image classification. `ImageDatagenerator`, shear range, rotation range are performed to transform the data into machine understandable form. Xception model architecture used to create the model. These input parameters are trained well in order to obtain a high accuracy. The features and quality of labeling process of the data utilized during the training period quantify the correctness of the outcome of the model.

4.2. *Data visualization and correlation matrix*



**Figure 3. Data Flow Diagram**

Correlation matrix is formed which establishes the relationship between each factor. It is formed used `Mathplotlib` library function. The matrix formed after applying linear regression to the data. Pickling is done to serialize the data and transport to the network. From `sklearn.ensemble RandomForestClassifier` is imported. The prediction model thus built is used to envisage the type of the cardiac disease the person is suffering from. Random Forest classifier gives the maximum accuracy.

4.3. Classification and prediction

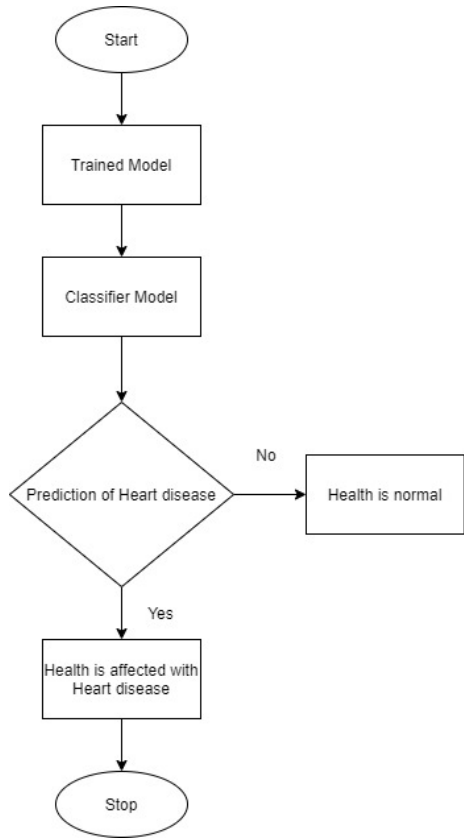


Figure 4. Proposed System Working Procedure

The noticed values are utilized to get some conclusions in this module. If inputs are given in this module, the values of results are foreseen. Labels of the results are applied to a dataset. The trained model is then used for classification based on the input parameters, this module will generate a value, based on this value, and the heart disease is being predicted.

5. Performance Analysis

The prevalent datasets were partitioned in 90:10 ratio for training and testing the model at each stage. After evaluation, as convolutional neural networks are the best for image classification, the accuracy obtained for CNN was 90% and the classification model built using random forest algorithm was obtained an accuracy of 87%. The Linear regression used to train the model using dataset has a success rate of 83%. The proposed model obtained a highest accuracy of 88%.

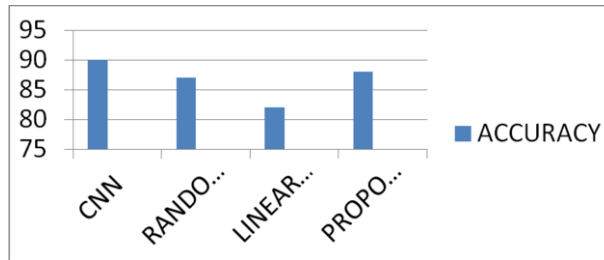


Figure 5. Comparison Graph of the Proposed Model

## 6. Conclusion And Future Works

There is an alarming increase in the number of cardiac deaths globally. Analysis and utilization of the prevalent raw data can help us build technological advanced systems which will be needful in times of an emergency. As, we know the early detection is the best prevention. This paper proposes a model to predict and classify the heart diseases by getting the chest x-ray and details of the patient. In comparison with the previous models, the proposed model has shown a significant improvement in the accuracy and obtained an accuracy 87%. If the proposed model could be implemented in large scale, we can reduce the uncertainty in the heart failures. In future, the prescription of medicines can also be included in the report based on the disease and more factors such as Electrocardiography can be added.

## References

- [1] Tamal MA, Islam MS, Ahmmed MJ, Aziz MA, Miah P, Rezaul KM. Heart disease prediction based on external factors: A machine learning approach. *International Journal of Advanced Computer Science and Applications (IJACSA)*. 2019;10(12):446-51.
- [2] Jagtap A, Malewadkar P, Baswat O, Rambade H. Heart disease prediction using machine learning. *International Journal of Research in Engineering, Science and Management*. 2019;2(2):352-5.
- [3] Raza A, Mehmood A, Ullah S, Ahmad M, Choi GS, On BW. Heartbeat sound signal classification using deep learning. *Sensors*. 2019 Jan;19(4819):1-15.
- [4] Patel J, TejalUpadhyay D, Patel S. Heart disease prediction using machine learning and data mining technique. *International Journal of Computer Science & Communication*. 2015 Sep;7(1):129-37.
- [5] Ajay H, Rao AR, Balavanan M, Lalit R. A Novel Cardiac Arrest Alerting System using IoT. *International Journal of Science Technology & Engineering*. 2017;3(10):78-83.
- [6] Dangare CS, Apte SS. Improved study of heart disease prediction system using data mining classification techniques. *International Journal of Computer Applications*. 2012 Jun;47(10):44-8.
- [7] Bhanumathi M, Priyanka JP, Murali Krishna J, Namrata. A Cardinal Approach for crack detection in a bridge using Automatic Pixel-level Technique. *International Journal of Grid and distributed Computing*. 2020;13(1):1011-7.
- [8] Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*. 2019 Jun 19;7:81542-54.
- [9] Ramprakash P, Sarumathi R, Mowriya R, Nithyavishnupriya S. Heart Disease Prediction Using Deep Neural Network. In 2020 International Conference on Inventive Computation Technologies (ICICT) 2020 Feb 26 (pp. 666-670). IEEE.
- [10] Alotaibi FS. Implementation of machine learning model to predict heart failure disease. *International Journal of Advanced Computer Science and Applications*. 2019;10(6):261-8.
- [11] Golande A, Pavan Kumar T. Heart disease prediction using effective machine learning techniques. *International Journal of Recent Technology and Engineering (IJRTE)*. 2019 Jun;8(1):944-50.