CNN-CardioAssistant: Deep Convolutional Neural Network and Recursive Feature Elimination Method for Heart Disease Detection

FARMAN HASSAN
University of Engineering and Technology Taxila
Auliya Ur Rahman
University of Engineering and Technology
Ali Javed (ali.javed@uettaxila.edu.pk)
University of Engineering and Technology Taxila
Ali Alhazmi
Jazan University
Majed Alhazmi
Jazan University

Research Article

Keywords: Cardiovascular disorder, coronary artery disease, smart Healthcare, Heart disease risk factors, Convolutional Neural Network, Statlog, Public Health dataset, Framingham dataset, Cleveland dataset, Z-Alizadeh sani dataset.

Posted Date: June 21st, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1721779/v1

Abstract

In recent times, we have seen an exponential rise in different chronic diseases due to our unhealthy lifestyles. Cardio disease is the most common and life-threatening among all diseases, which contributes to a very high mortality rate. Accurate detection of cardio disease at an early stage is vital to save the lives of people. Most of the existing cardiovascular disease detection systems suffer from lower performance and efficiency due to redundant attributes, dimensionality curse, imbalance, and noisy datasets. In this work, we proposed a novel convolutional neural networks-based system (CNN-cardioAssistant) that predicts cardiovascular disease in patients. Recursive feature selection (RFE) is employed to select more prominent features from the clinical data of cardio patients. The selected features are then used to train the proposed CNN-CardioAssistant as well as 11 different classifiers i.e., support vector machine (SVM), Random Forest, Decision Tree, logistic regression, Naïve Bayes, K-Nearest Neighbor, XGboost, Multi-layer Perceptron, Gaussian process classifier, Adaboost, and Quadratic discriminant analysis separately for cardio disease prediction. We compared the results of all the methods on three subsets of features i.e., 6,8, and 15 for each dataset. The features selection method provides optimal subsets of features that can reliably be used to predict cardiovascular disease with the highest accuracy. Experimental results on three different cardio datasets i.e., Public Health, Framingham, and Z-Alizadeh sani clearly demonstrate that the proposed CNN-CardioAssistant system has superior performance against the existing state-of-the-art methods.

1. Introduction

Cardiovascular disorder is a life-threatening disease that affects the normal action of the heart. Cardiovascular disorder occurs if the person's heart is not able to pump sufficient blood supply to satisfy the human body's needs. There are two reasons for heart disease such as blockage or narrowing the coronary arteries since arteries deliver blood to the heart. The diagnosis of cardiovascular disease is mostly based on symptoms and physical examination. Research community has found some major risk factors i.e., obesity, usage of alcohol, family history of heart diseases, lack of physical exercise, smoking, high blood pressure, diabetes mellitus, high blood cholesterol, mental stress, sleep apnoea, etc., which can increase blood vessels and coronary heart diseases [3]. According to the world health organization (WHO) [1], an estimated 17.7 million people died from heart diseases resulting in 31% of deaths around the world, and by 2030 around 23.6 million people may die from cardiovascular disease. It has also been observed that 25% people usually die suddenly from coronary arteries diseases without any prior symptoms [2]. Clinical physicians treat heart patients but there is a shortage of cardiologists, and they often experience difficulties in accurately predicting the presence of cardiovascular disease in patients. More specifically, manual detection of cardiovascular disease is a less accurate, costly, and time-consuming task. People also ignore heart disease until it becomes severe due to various reasons i.e., expensive treatment, etc. Research community is working hard to design intelligent models for timely and effective detection of cardiovascular disease. Intelligent clinical decision-making-based systems have been designed to help physicians for accurate diagnosis of cardio diseases in a short period [6]. Healthcare industries collect a maximum amount of clinical data that contains useful information of patients, which can be used for effective prediction and providing an appropriate result. These results can be used to prevent and reduce the cost of expensive tests and surgical treatments.

Research community explored the effectiveness of machine learning methods i.e., support vector machine (SVM), decision tree (DT), random forest (RF), logistic regression (LR), naïve bayes (NB), multi-layer perceptron (MLP), quadratic discriminant analysis (QDA), etc., which obtained good results in the classification problems over other data classification models [4]. SVM and RF are the most commonly used classifiers for binary classification problems [5]. Research shows that SVM performs better over other statistical algorithms as well as instance-based learning methods [7]. RF is an ensemble of unpruned DT and is employed in case a of large number of features and training datasets. RF comprises of many DTs and produces the output class of an individual DT [11]. DT is an analytical method, which can be employed to represent the classification problems. Additionally, DTs are helpful in the prediction of marketing, economics, engineering, and medicine [10]. NB is effective [9] for text classifications, system performance management as well as medical diagnosis applications. KNN technique is employed to categorize the samples in feature space-based on nearby samples. KNN is a supervised machine learning algorithm that is used for regression and classification problems. KNN is an instance-based learning method for assuming instances related to points in an n-dimensional Euclidean space and discrete-valued target functions. The target function value for fresh query is merely estimated locally and overall, the calculations are changed until the final classification [12]. Recently, the research community proposed numerous machine learning techniques [13–21] in diagnosing various diseases. Classification in medical sciences is the most important and popular decision-making tool for the diagnosis of diseases.

Research community explored various cardiovascular disease prediction systems [24–29], [44–48], [51, 54] based on machine learning algorithms. In [24], different models were employed for the prediction of heart diseases i.e., approximated L0-norm SVM strategies, recursive characteristic elimination, and standard SVM, but able to achieve an accuracy of 76.8% only due to using unnecessary features. In [25], SVM was employed for the detection of cardiovascular disease in patients. Removing the unnecessary features improved the accuracy of this method to 85%. In [26], the system employed a set of strategies and multivariate adaptive regression splines for the selection of features in order to decrease the size of explanatory features. The system achieved an accuracy of 82.14% using LR as a classifier. In [27], various machine learning algorithms i.e., KNN, MLP, SVM were used with particle swarm optimization (PSO) for the detection of cardiovascular disease. PSO used two types of optimizations i.e., relative, and brief optimization. This system obtained an accuracy of 81.73% on KNN, 82.30% on MLP, and 75.37% on SVM. In [28], ant colony optimization (ACO) was merged with the KNN to detect heart disease. This method achieved an accuracy of 70.26% using the combination of ACO and KNN. In [29], machine learning classifiers i.e., SVM, NB, and DT-GI were used to detect cardiovascular disease. In [44], NuSVM called N2Genetic-NuSVM and two-level hybrid genetic algorithm was used to classify the healthy persons and cardiovascular patients. An accuracy of 93.08% was yielded using a 10-fold cross-validation technique. In [45], numerous data mining models i.e., SVM, random trees (RTs), Chi-squared automatic interaction detection (CHAID), and C5.0 were employed using 10-fold cross-validation. RTs performed well and yielded an accuracy of 91.47% on 10-fold cross-validation while the C5.0, SVM, and CHAID achieved the accuracy of 82.17%, 69.77%, and 80.62%, respectively. In [46], a classification and regression tree (CART) was used with 40 clinical features to develop a heart disease diagnosis system. In [48], a heterogeneous features selection algorithm (2HFS) was developed for selecting important features and employed machine learning techniques i.e., DT, NB, RF, and XGboost to detect the heart disease. To handle the imbalance in dataset, two methods i.e., adaptive synthetic and synthetic minority over-sampling technique (SMOTE) were used and achieved an accuracy of 92.58%. In [51], heart disease prediction system was designed using large amount of medical data available on the cloud. SVM was employed to classify between the healthy and cardio patient. The cloud-based recommendation system used history of patients, streaming medical data, and knowledge database to provide users alert and recommendations. In [54], hybrid random forest linear model (HRFLM) was proposed to detect the cardiovascular disorder in patients and employed different machine learning methods i.e., SVM, NB, NN, and RF. HRFLM achieved the best results (accuracy of 88.7%) among all the machine learning methods.

Research community also explored deep learning methods [22, 23]-[30–34] for the detection of cardiovascular disease. Deep learning (DL) methods are growing rapidly and imitate the way humans gain certain types of knowledge. DL represents multi-level records and is able to overcome the selectivity invariance problem [8]. These methods are used in various application domains, which work with a large amount of data and have the capability to decode the complicated hassle easily. In [22], a smart health care system was proposed for cardiovascular disease prediction using sensor data and electronic medical records. Ensemble deep learning was employed for the classification purpose and yielded an accuracy of 98.5%. In [23], an optimized deep neural network was employed to detect the presence of cardiovascular disease. This method used the Talos optimization technique and achieved an accuracy of 90.78%. In [30], neural network and genetic algorithm were used to make a hybrid system for the classification of heart disease. The weights of each neuron were initialized with the help of a global optimization technique and genetic algorithm that achieved an accuracy of 89%. In [31], cardiovascular disease prediction system named CardioHelp was designed for the detection of heart disease. Convolution neural network was employed to classify between the cardiovascular disease patient and healthy person, which achieved an accuracy of 97%. In [32], neural network (NN) was used for the detection of heart disease and achieved an accuracy of 96.30% on three-fold cross validation. In [33], multiple layer neural network was used for cardiovascular disease prediction. Two hidden layers between an input and output were used. An average accuracy of 91.60% was achieved using multiple layer neural network. In [34], two classification strategies i.e., locally linear embedding and regression were used for the classification purpose of coronary artery disease and achieved an accuracy of 80%. In [53], an ensemble learning system was proposed using the LR, NB, and MLP for diagnosing the heart disease. Majority voting role was employed to combine the output of all the three classifiers. The ensemble learning system yielded an accuracy of 88.88%. In [49], an IoT based heart disease detection system was proposed using the modified deep convolutional neural network for classification purpose. Two devices i.e., heart monitor device and smart watch were attached to the patient for checking blood pressure and electrocardiogram. An accuracy of 98.2% was achieved using mapping-based cuttlefish optimization algorithm for features selection. In [47], four distinctive features selection methods i.e., information gain, gini index, principal component analysis, and weights by SVM were used to select optimal subset of features whereas, four optimization techniques i.e., evaluation strategy, particle swarm optimization, forward and backward weights optimization were employed to enhance the neural network using 10-fold cross validation technique.

Existing methods also explored applications [35, 36, 37] for prediction of heart disease. In [35], an application using machine learning algorithms were used for monitoring of the cardio patients. A cloud-based system was presented to provide an interface to the users for uploading physiological data to check the status of their cardiac health. The application achieved an accuracy of 77.39%, 95.05%, 95.76%, 97.53% and 86.40% on ANN, LR, RF, SVM and NB, respectively. In [36], cardiovascular disease prediction system based on the fuzzy cognitive map and structural equation modeling using Canadian community health survey dataset was used. There are 20 attributes used in this dataset for validation purpose. The system achieved an accuracy of 74%. In [37], an intelligent cardiovascular disease prediction system was introduced using the big data. The system uses data mining techniques, NB, and a software framework Hadoop and Mahout for classification purposes.

Even, if the literature signifies much improvement to detect the heart disease, the problem is far being solved. Still, there exists certain limitations of the existing state-of-the-art heart attack prediction methods [44–57] i.e., low accuracy, expensive treatment, unable to better handle missing values and high dimensional datasets, undiscovered valuable information from clinical datasets, usage of irrelevant variables which degrades the system performance, using small datasets for training purpose that are unreliable for heart disease prediction, and incapable of extracting prominent features from diverse datasets. Therefore, to achieve the maximum capability of machine learning and deep learning algorithms, it is necessary to pre-process the clinical data to achieve better classification accuracy because clinical datasets are irregular and uncertain. So, we need to develop an effective and robust system, which practitioners can use to reduce the death rate of those patients having the cardiovascular disorders. The major contributions of our work are as under:

- We present an effective features selection method by using the recursive feature elimination (RFE) technique to identify an optimal subset of prominent features.
- We proposed a novel convolutional neural network CNN-cardioAssistant to detect cardiovascular disease patients.
- We present a reliable and efficient prediction model using the RFE and CNN-cardioAssistant to effectively diagnose the cardiovascular disease.
- We performed rigorous experiments on three diverse datasets as well as on cross dataset settings to check the generalizability and robustness of the proposed system over state-of-the-art methods.

The remaining paper provides the details of the proposed method, experimental setup & results, and the conclusion of our work.

2. Proposed Methodology

This section presents the detailed working mechanism of the proposed method. The major objective of the proposed system is to predict the presence of cardiovascular disease in patients at preliminary stages based on past records. The workflow of the proposed system is comprised of three phases i.e., features selection, classification, and validation analysis. The initial phase deals with the procedure of important features selection by employing the RFE. In the second phase, optimal subsets of features are used to train CNN-CardioAssistant for the prediction of cardiovascular disease. Finally, in the third phase, the performance of the system is validated. The proposed system takes clinical data as input comprising of age, sex, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, resting electrocardiographic results, maximum heart rate achieved, exercise induced angina, ST depression induced by exercise relative to rest, the slope of peak exercise ST segment, number of major vessels colored by fluoroscopy, thal (3 = normal, 6 = fixed defect, 7 = reversible defect), and angiographic disease status. The proposed working mechanism is shown in Figure I.

2.1. Pre-Processing

Pre-processing is an important stage to refine the dataset. It is noteworthy to mention that in most cases, the missing values and redundant features degrades the accuracy of the method. So, we initially check the dataset for the missing values and redundant features. We can handle missing values in many ways i.e., ignore all values, replace the values with some numeric type values, replace it with frequently appearing value for that feature, replace it with mean value, etc. In Statlog, and Public Health dataset is no missing values, but in Cleveland dataset, there are some missing values. The missing values in Cleveland dataset are the "value of main vessels" and the "thal" features. The pre-processing stage comprises of two steps i.e., removing the redundant features and replacing the missing values. We observed that three attributes such as patient's blood pressure, cholesterol, and age group have missing values. We replaced the missing values of these specific attributes by computing the mean value of the column.

2.2. Features Selection

Feature selection is an important process to enhance the effectiveness of classification performance. Hence, selecting a subset of more reliable features is crucial to make the challenging prediction task easy. The feature selection process can be performed in numerous ways, i.e., filtering the data, grouping the data to make clusters, eliminating irrelevant features by applying feature selection algorithms, etc. The best subset of features can be obtained by evaluating the features on some criteria i.e., dependency measures, distance measures, consistency measures, and information measures. Any dataset comprises of three distinct types of features i.e., relevant, irrelevant, and redundant. Relevant features hold maximum information about the data, irrelevant features don't hold any valuable information, and the redundant features hold information, which is already given via other features in the dataset and hence, the redundant features don't contribute to predict the salient data patterns. So, in this paper, we employed the RFE selection method [38] to eliminate the irrelevant and redundant features and finally obtain the most discriminative and reliable features, which improves the accuracy of the model. RFE is a cross-validated features selection method that recursively selects a subset of features based on the ranking of features. RFE removes the collinearity and dependency among the features of the dataset. RFE eliminates the features and recursively builds the model with the remaining feature until all the features of the dataset are finished. At each iteration, features rank is measured and features with the lowest rank are eliminated. The features selection algorithm selects first *n* features from the ranking. The process is given below.

- Build the classification model using an optimal subset of features.
- Compute features rank to check prominent features of dataset.
- Eliminate features with low rank.

The pseudo code of RFE method is presented in Algorithm 1.

Algorithm 1

Inputs:

Training set S

Set of n features $K = \{f_1, \ldots\ldots, f_n\}$

Ranking technique $R\,(S, K)$

Outputs:

Final Ranking F_r

Code:

Repeat for x in $\{1{:}\ n\}$

Rank optimal subset using $R\,(S, K)$

$f^* \longleftarrow$ last ranked feature in K

$F_r\,(n\text{-}x+1) \longleftarrow f^*$

$K \longleftarrow K - f^*$

2.3. Convolutional Neural Networks

The prediction of cardiovascular disease is a binary classification problem. Neural networks are effective classifiers and have shown significant enhancements under certain scenarios [58]. Research community has explored neural networks by using different configurations with application-specific settings such as various hidden layers, neurons, and minibatch sizes to improve the classification accuracy. Application areas in which results are improved by employing neural networks include speech processing, time series prediction, and image processing [59]. Artificial neural networks perform the transformation of the input data over hidden layers and estimate errors at output layers [60–62]. After transformation of input data and normalization, the gradient descent algorithm utilizes backpropagated error by an output layer to iteratively update the weights of the layers. The gradient descent algorithm has been enhanced by using rigorous experimentation and analysis such as making modifications to the algorithm, visualization of hidden layers, reducing the overfitting problems, scheduling the training process, and making the neural network layers nonlinear. However, the problem is far from being solved despite significant improvements in neural network applications. The training networks overfit due to irrelevant parameters in deep learning architecture and the problems get worse when the training data is insufficient for training the network. Numerous data augmentation techniques [63, 64] such as position augmentation and color augmentation have been proposed to address the issues of overfitting the models. Data augmentation artificially generates data from the existing data by employing various settings such as cropping, adding noise, shear, translation, etc., however, these techniques fail miserably on clinical datasets. A small amount of training data leads to poor training that results in an inaccurate and poor classification. As it is obvious that wrong decision in the medical field has great consequences and accompanies bigger penalty compared to other applications such as image synthesis, speech processing, semantic labeling, chat-bot configuration, etc.

Due to an inaccurate detection mechanism, a person having cardiovascular disease may be left untreated, which can lead to incorrect therapeutic medication as an accurate prediction in medical applications are necessary, therefore, our main goal is to enhance the classification accuracy and precise prediction. We proposed a CNN-based model to effectively detect cardiovascular patients. We also compared the performance of the traditional machine learning algorithms with the proposed system and performed cross dataset experiments to check the generalizability of the system. The detailed architecture of the proposed CNN-based model is discussed in the subsequent section.

2.4. Proposed CNN-cardioAssistant Architecture

In order to enhance the classification accuracy and precise prediction of cardiovascular disease. We proposed a CNN-based model to effectively detect cardiovascular patients and named our system as the Convolutional neural networks-based cardiovascular assistant (CNN-CardioAssistant). We also compared the performance of the traditional machine learning algorithms with the proposed system and performed cross dataset experiments to check the generalizability of the system. The detailed architecture of the proposed CNN-based model is discussed in the subsequent section.

The detailed proposed CNN-based model is shown in Figure II. The architecture of the proposed CNN-based model is a feedforward network, which works on a sequential single input and single output way. In this work, we performed binary classification experimentations and assumed that patients having the cardiovascular disease were classified as “1” while the absence of cardiovascular disease is classified as “0”. The number of attributes selected by employing RFE is 6, 8, and 15. We split the data into 80/20 and used 80% data for training while 20% for testing the trained model. More specifically, we used 240 records of the Z-Ali Zadeh Sani dataset for training the model while 60 records for testing the trained model. From the Framingham dataset, we used 3,390 records for training while 848 records for testing purposes. Similarly, we used 820 records of the Public Health dataset for training while 105 records for evaluating the model. Therefore, the input layer as shown in Figure II. has $R^{240 \times 6}$, $R^{240 \times 8}$, and $R^{240 \times 15}$ dimensions for Z-Ali Zadeh Sani dataset while $R^{240 \times 6}$, $R^{240 \times 8}$, and $R^{240 \times 15}$ dimensions for Framingham dataset and $R^{240 \times 6}$, $R^{240 \times 8}$, and $R^{240 \times 13}$ dimensions for the public Health dataset.

The optimal subsets of features by RFE are combined in a fully connected Dense layer with 32 neurons. The Dense layer efficiently normalizes numerous attributes before the non-linear transformation by employing the rectified linear unit. We also used the dropout layer equal to 0.5 to reduce the overfitting problem. We used a cascaded set of four convolutional layers that are followed by dense layers. In the first convolutional layer, we used hidden units size of 256, kernel size of 3, strides of 1, padding equal to the same, and an activation function equal to Relu. The first convolutional layer is followed by a batch normalization layer, which is then followed by a max pooling layer of pool size equal to 3, strides equal to 1, and padding equal to the same. The 2nd and 3rd have the same configurations while we have used different configurations for the last fourth convolutional layer. The fourth layer has hidden units of size 64, kernel size of 3, strides of 1, padding equal to the same, and an activation layer of Relu. The fourth layer is followed by the dense layer that has an activation function of relu. We also used rmsprop as an optimizer, learning rate equal to 0.005, decay equal to 0, and set the loss function as categorical entropy loss. We used the above-mentioned hyperparameters to achieve better performance classification precision and accuracy for cardiovascular disease prediction.

3. Experimental Results And Discussion

This section provides the details of the experiments conducted to measure the performance of the proposed system to detect cardiovascular disease. We evaluated the performance of the proposed system using the accuracy, precision, recall, and F1-score. The details of experiments on the datasets are given below.

3.1. Datasets

In this work, we used five popular cardiovascular disease datasets i.e., Statlog [39], Cleveland [40], Public Health [41], Z-Alizadeh sani [42], and Framingham datasets [43] that are publicly available in UCI machine learning repository database. Each dataset has different attributes such as Statlog has an age, gender, chest pain type, resting blood pressure, serum cholesterol, fasting blood sugar, maximum heart rate, exercise-induced angina, oldpeak, the slope of the peak, number of major vessels, and thal, etc., while the Cleveland and public health datasets have also the same attributes as that of Statlog dataset. Z-Alizadeh Sani Dataset has a total of 54 features and is arranged in four groups such as demographic, symptom and examination, ECG, and laboratory and echo features. Framingham dataset has 4238 records, which belong to three groups such as demographic, behavioral, and medical risk factors. This dataset provides a potential risk of coronary artery disease before 10 years. Each dataset contains the data of both male and female patients. The datasets are diverse in terms of attributes where each dataset has distinctive features from others. The details of all the five datasets are given in Table I.

Table I. Details of the datasets.

Dataset	No of Observations	No of attributes	No of healthy persons	No of heart patients
Statlog	270	75	150	120
Cleveland	303	76	164	139
Z-Alizadeh sani	300	54	87	216
Framingham	4238	16	3596	644
Public Health	1025	14	499	526

3.2. Results on Z-Alizadeh sani Dataset

The objective of this experiment is to check the effectiveness of the proposed system on the Z-Alizadeh sani dataset [42] to detect the presence of cardiovascular disease using three different subsets of features i.e., 6, 8, and 15. We employed RFE feature selection technique to select an optimal subset of features that contain the maximum information of the cardiovascular disease to train the 11 machine learning algorithms and the proposed CNN-cardioAssistant. We conducted experiments on all the three optimal subsets of features.

In the first phase, we employed RFE and selected an optimal subset of six features i.e., age, body mass index (BMI), typical chest pain (ca), triglyceride (TG), platelet (PLT), and ejection fraction (EF-TTE) to train multiple classifiers for cardiovascular disease prediction. We achieved an accuracy of 80.32%, 81.96%, 85.24%, 85.24%, 83.60%,67.21%, 83.60%, 78.68%, 81.96%, 85.24%, 80.12%, and 88.52% on DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. From Table II, we can observe that the proposed method CNN-cardioAssistant performed well and attained maximum accuracy of 88.52%, precision of 100%, recall of 72.13%, and F1-score of 83.81% on an optimal subset of six features among all the twelve methods. KNN, XGboost, and RF performed second best and achieved an accuracy of 85.24%, precision of 90.90%, recall of 80.80%, and F1-score of 89.89% while the GPC performed the worst by achieving an accuracy of 67.21%, precision of 72.72%, recall of 80% and F1-score of 76.19%. From the results on the optimal subset of six features, we can conclude that the combination of these features does not hold enough information and can't reliably be used by the cardiologists for the accurate detection of cardiovascular disease. So, we need to include more prominent features to further enhance the performance of our system.

In the second phase, we select another optimal subset of features comprised of two additional features i.e., fasting blood sugar (FBS) and erythrocyte sedimentation rate (ESR) along with the previous six selected features. This subset of eight features includes the following i.e., age, BMI, typical chest pain, FBS, TG, ESR, PLT, and EF-TTE. We achieved an accuracy of 81.96%, 81.96%, 81.96%, 93.44%, 78.68%, 60.65%, 85.24%, 80.32%, 80.32%, 86.88%, 81.30%, and 88.52% on DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. We achieved the best accuracy of 93.44%, precision of 95.45%, recall of 89.36%, and F1-score of 92.30% on these eight selected features using the (RFE_XGboost) method. Moreover, the proposed system also performed second best and achieved an accuracy of 88.52%, precision of 100%, recall of 72.13%, and F1-score of 83.80%. (RFE_GPC) performed worst by achieving an accuracy of 60.65%, precision of 75%, recall of 75%, and F1-score of 75%. From the results on the subset of eight features, we observed that by adding the additional two features, the accuracy of each method increases and found significant improvement for XGboost where an increase of 8.20% was observed.

In the third phase, we increased the number of input features from eight to fifteen i.e., age, weight, BMI, blood pressure (BP), ca, FBS, TG, low-density lipoprotein (LDL), ESR, hemoglobin (HB), Na, white blood cells (WBC), Lymph, PLT, and EF-TTE. We achieved an accuracy of 85.24%, 78.68%, 67.21%, 80.32%, 72.13%, 24.59%, 81.96%, 86.88%, 81.96%, 83.60%, 78.60%, and 78% on DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. The detailed results of all the three subsets of features in terms of accuracy, precision, recall, and F1-score are reported in Table II. We achieved the best accuracy of 86.88%, precision of 90.91%, recall of 90.91%, and F1-score of 90.90% on (RFE_NB) method. The DT performed second best and achieved an accuracy of 85.24%, precision of 95.45%, recall of 85.71%, and F1-score of 90.30% while the GPC performed the worst and achieved an accuracy of 24.59%, precision of 0.00%, recall of 0.00%, and F1-score of 0.00%. We can conclude from the results that (RFE_XGboost) performed well on an optimal subset of eight features among all different methods and subsets of features. For this dataset, the combination of ideal features is eight for (RFE_XGboost). The addition of two features i.e., FBS and ESR to a subset of 6 features played a significant role in the correct classification of healthy people and cardiovascular patients. It can be concluded that (RFE_XGboost) on eight features can reliably be used in clinics by cardiologists for the detection of cardiovascular disease.

Table II. Evaluation on Z-Alizadeh sani dataset using machine learning techniques on 6, 8 and 15 features subsets.

Algo	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precisi
DT	6	80.32	90.90	83.33	86.96	8	81.96	95.45	83.33	89.00	15	85.24	95.45
LR		81.96	90.90	85.10	87.91		81.96	93.18	85.41	89.10		78.68	93.18
KNN		85.24	90.90	88.88	89.89		81.96	93.18	78.84	85.40		67.21	84.09
XGboost		85.24	90.90	88.88	89.89		**93.44**	**95.45**	**89.36**	**92.30**		80.32	93.18
MLP		83.60	97.72	82.69	89.58		78.68	93.18	83.67	88.20		72.13	100
GPC		67.21	72.72	80.00	76.19		60.65	75.00	75.00	75.00		24.59	0.00
AB		83.60	93.18	85.41	89.13		85.24	97.72	86.00	91.50		81.96	93.18
NB		78.68	85.71	83.72	84.71		80.32	86.36	86.36	86.40		**86.88**	**90.91**
QDA		81.96	86.6	88.37	87.36		80.32	86.36	86.36	86.40		81.96	90.91
RF		85.24	93.18	87.23	93.18		86.88	97.72	87.75	92.50		83.60	95.45
SVM		80.12	90.10	83.12	86.70		81.30	93.12	82.22	88.70		78.60	93.10
Proposed		**88.52**	**100**	**72.13**	**83.81**		88.52	100	72.13	83.80		78.00	66.67

3.3. Results on Framingham Dataset

This experiment is designed to check the effectiveness of distinctive features combination of the proposed system on the Framingham dataset [43] to accurately detect cardiovascular disease. As earlier, we conducted experiments on all features of the dataset as well as on two subsets of optimal features i.e., a combination of six and eight features.

In the first stage, we employed a subset of six features i.e., age, total cholesterol level (totchol), systolic blood pressure (sysBP), diastolic blood pressure (diaBP), BMI, and blood glucose level (glucose) to train multiple classifiers independently for classification. We achieved an accuracy of 83.96%, 85.25%, 83.37%, 85.02%, 84.78%, 75.82%, 85.02%, 82.19%, 82.66%, 84.66%, 82.60%, and 77.30% on DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. From the results reported in Table III, we can observe that (RFE_LR) outperformed among the twelve methods on a combination of six features and achieved an accuracy of 85.25%, precision of 5.55%, recall of 29,16% and F1-score of 9.30%. (RFE_AB) also achieved an accuracy of 85.02%, precision of 7.14%, recall of 47.36%, and F1-score of 12.40% while we achieved the lowest performance on (RFE_GPC) and achieved an accuracy of 75.82%, precision of 16.66%, recall of 17.35%, F1-score of 17.00%. We observed that all twelve methods provide lower performance in terms of precision, recall and F1-score. These six features do not contain enough details, so, we need to use more features in order to improve the performance of the system.

In the second stage of this experiment, we added two more features i.e., heartrate and cigsPerDay to a subset of six features. The second optimal subset of features is comprised of eight features i.e., age, cigrate person smoke per day (cigsPerDay), totchol, sysBP, diBP, BMI, heartrate, and glucose. We achieved an accuracy of 83.25%, 85.25%, 84.19%, 84.90%, 85.02%, 76.88%, 84.78%, 82.31%, 81.95%, 84.90%, 81.80%, and 83.24% for DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. We achieved best accuracy of 85.25% on (RFE_LR), precision of 2.38%, recall of 60.00%, F1-score of 4.58%. (RFE_MLP) performed second best and achieved an accuracy of 85.02%, precision of 12.69%, recall of 48.48%, and F1-score of 20.13% while the (RFE_GPC) performed worst and achieved an accuracy of 76.88%, precision of 20.63%, recall of 21.31%, and F1-score of 20.97%. From the results reported on a subset of eight features, we observed that all the methods again performed worst in terms of precision, recall and F1-score. This combination of features is not reliable to be used for the prediction of cardiovascular disease. We also observed that adding these two features i.e., heart rate and cigsPerDay enhance the accuracy on CNN-cardioAssistant from 77.30–83.24% while the accuracy of other methods increases slightly.

Table III. Evaluation on Framingham dataset using machine learning techniques on 6, 8, and 15 features subsets.

Algo	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precisi
DT	6	83.96	5.55	29.16	9.30	8	83.25	3.96	19.23	6.57	15	85.24	2.38
LR		85.25	2.38	60.00	4.60		85.25	2.38	60.00	4.58		88.52	5.55
KNN		83.37	7.14	6.76	6.90		84.19	9.52	37.50	15.19		70.49	9.52
XGboost		85.02	3.17	44.44	5.90		84.90	5.55	43.75	9.85		86.88	5.55
MLP		84.78	4.76	40.00	8.50		85.02	12.69	48.48	20.13		78.68	23.60
GPC		75.82	16.66	17.35	17.00		76.88	20.63	21.31	20.97		37.70	20.63
AB		85.02	7.14	47.36	12.40		84.78	8.73	44.00	14.57		90.16	21.38
NB		82.19	23.01	34.94	27.80		82.31	24.60	36.04	29.25		86.88	2.38
QDA		82.66	18.25	34.32	23.80		81.95	20.63	32.91	25.37		81.96	21.64
RF		84.66	5.55	38.88	9.70		84.90	5.55	43.75	9.85		84.90	5.55
SVM		82.60	18.10	34.80	27.50		81.80	20.50	32.80	25.20		85.88	2.38
Proposed		**77.30**	**18.18**	**14.29**	**16.00**		**83.24**	**28.57**	**7.14**	**11.43**		**99.86**	**100.00**

In the third stage of this experiment, we increased the number of input features from eight to fifteen. This optimal subset of features consists of gender, age, education, currentSomker, cigsPerDay, BPMeds, prevalentStroke, prevalentHyp, diabetes, totChol, sysBP, diaBP, BMI, heartrate, and glucose. We achieved an accuracy of 85.24%, 88.52%, 70.49%, 86.88%, 78.68%, 37.70%, 90.16%, 86.88%, 81.96%, 84.90, 85.88%, and 99.86% for DT, LR, KNN, XGboost, MLP, GPC, AB, NB, QDA, RF, SVM, and CNN-cardioAssistant, respectively. The detailed results of all methods on all the three subsets of features in terms of accuracy, precision, recall, and F1-score are given in Table III. We observed that our method CNN-cardioAssistant performed well among all other methods and achieved an accuracy of 99.86%, precision of 100.00%, recall of 99.11%, F1-score of 99.55%. (RFE_AB) performed second-best on this subset of optimal features and achieved an accuracy of 90.16%, precision of 21.38%, recall 60.00%, and F1-score of 15.58% while the GPC performed the worst and achieved an accuracy of 37.70%, precision of 20.63, recall of 60.00%, and F1-score of 20.97%. From the results, it is concluded that increasing the number of features enhances the accuracy of the system. Comparing the results of eight and fifteen features, we noticed a slight increase in accuracy on all the classifiers except the CNN-cardioAssistant where we experienced a significant improvement from 83.24–99.86%. From the detailed results, we can conclude that CNN-cardioAssistant outperformed against all the other methods. The proposed method using the combination of fifteen features can reliably be used by the clinical physicians and medical specialists in hospitals, health care, and medical centers for the prediction of cardiovascular disease patients early and accurately.

3.4. Results on Public Health Dataset

The objective of this experiment is to check the effectiveness of the proposed system on the Public Health dataset [41] to classify healthy persons and cardiovascular disease patients. For this purpose, we designed a three-phase experiment on the subset of six, eight, and thirteen features.

In the first phase, we employed RFE to select an optimal subset of six features i.e., age, cp, chol, thalach, pldpeak, and ca. We achieved an accuracy of 90.73%, 83.41%, 77.56%, 92.68%, 83.41%, 87.31%, 83.90%, 84.87%, 83.31%, and 98.04% on DT, LR, KNN, XGboost, MLP, AB, NB, QDA, SVM, and CNN-cardioAssistant, respectively. From Table IV, we observe that in the first experiment on an ideal subset of six features, CNN-cardioAssistant performed well among all other methods and achieved an accuracy of 98.04%, precision of 100%, recall of 96.39%, and F1-score of 98.17%. The technique (RFE_XGboost) performed second best and achieved an accuracy of 92.68%, precision of 97.19%, recall of 89.65%, and F1-score of 93.28% while the (RFE_KNN) performed worst and achieved an accuracy of 77.56%, precision of 74.76%, recall of 80.80%, and F1-score of 77.67%. We can conclude from these results that CNN-cardioAssistant performs well even on the reduced features-set of six and can still be used for the correct prediction of heart disease.

In the second phase, again, we employed RFE to select eight key features. This optimal subset of features has same six features that were previously selected while the other two prominent features such as trestbps and thal are added to it. We achieved an accuracy of 91.70%, 84.87%, 74.63%, 95.12%, 82.43%, 91.70%, 86.34%, 84.39%, 82.33%, and 98.04 on DT, LR, KNN, XGboost, MLP, AB, NB, QDA, SVM, and CNN-cardioAssistant, respectively. From the Table IV, we can observe that CNN-cardioAssistant outperformed again against all the methods and achieved an accuracy of 98.04%, precision of 100%, recall of 96.40%, and F1-score of 98.17%. The (RFE_XGboost) also achieved better accuracy of 95.12%, precision of 98.13%, recall of 92.92%, and F1-score of 95.45% while the performance of the (RFE_KNN) degrades and achieved an accuracy of 74.63%, precision of 73.83%, recall of 76.60%, and F1-score of 75.24%. From the results on subset of eight features, we observed that by adding the two features i.e., testbps and thal enhances the accuracy of other methods slightly but the accuracy of CNN-cardioAssistant remains the same. We didn't observe much increase in accuracy, so, we can conclude that both combinations of features i.e., six and eight are reliable for the accurate prediction of cardiovascular disease patients.

In the third phase, we increased the number of features from eight to a combination of thirteen features i.e., age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, and thal. We achieved an accuracy of 78.68%, 85.24%, 63.93%, 85.25%, 78.68%, 90.16%, 85.24%, 83.60%, 84.25%, and 98.68% on DT, LR, KNN, XGboost, MLP, AB, NB, QDA, SVM, and CNN-cardioAssistant, respectively. Again, we achieved maximum accuracy of 98.68% on CNN-cardioAssistant. The results on each subset in terms of accuracy, precision, recall, and F1 score are given in Table 4. Our proposed method CNN-cardioAssistant performed the best and achieved an accuracy of 98.68%, precision of 100%, recall of 97.18%, and F1- score of 98.57%. The (RFE_AB) also performed well and achieved an accuracy of 90.16%, precision of 88.23%, recall of 93.75%, and F1-score of 90.91% while the performance of (RFE_KNN) degrades and achieved an accuracy of 63.93%, precision of 64.70%, recall of 68.75%, and F1-score of 66.67%. The experimental results on thirteen features illustrate that there is a minor increase of 0.64% in an accuracy of CNN-cardioAssistant but we observe a decline in accuracy of other methods. We can conclude from the results on all the three combinations of features that CNN-cardioAssistant performed the best and can reliably be used for the timely and accurate prediction of cardiovascular disease. These selected features contain maximum information that is required to make an accurate prediction of cardiovascular disease.

Table IV. Evaluation on Public Health dataset using machine learning techniques on 6, 8, and 13 features subsets.

Algo	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precision%	Recall%	F1-Score%	No of Attr	Accuracy%	Precisi
DT	6	90.73	92.52	90.00	91.24	8	91.70	97.20	88.14	92.45	13	78.68	79.41
LR		83.41	85.98	82.88	84.11		84.87	87.85	83.93	85.85		85.24	87.87
KNN		77.56	74.76	80.80	77.67		74.63	73.83	76.60	75.24		63.93	64.70
XGboost		92.68	97.19	89.65	93.28		95.12	98.13	92.92	95.45		85.25	87.87
MLP		83.41	85.98	82.88	84.40		82.43	94.39	80.80	97.07		78.68	91.17
AB		87.31	88.78	87.15	87.97		91.70	94.39	90.18	92.24		90.16	88.23
NB		83.90	88.78	81.89	85.20		86.34	88.79	85.59	87.16		85.24	91.17
QDA		84.87	91.58	81.66	86.35		84.39	88.79	82.61	8559		83.60	85.29
SVM		83.31	85.80	82.70	83.90		82.33	94.29	80.69	96.80		84.25	86.91
Proposed		**98.04**	**100**	**96.39**	**98.17**		**98.04**	**100**	**96.40**	**98.17**		**98.68**	**100**

3.5. Performance comparison on cross datasets

This experiment is designed to evaluate the performance of the proposed system on the cross-datasets settings to check the robustness and generalizability of the proposed system. For this purpose, we analyzed the three datasets such as Cleveland, Statlog, and public health. We observed that these datasets have similar attributes, thus, we selected the thirteen similar attributes such as age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope, ca, and thal of these datasets for cross dataset experiments. More specifically, in first cross dataset scenario, we used Cleveland dataset for training the model and the public health dataset for the testing purpose. We achieved the highest accuracy of 98.44% on CNN-cardioAssistant. The detailed results in terms of accuracy, precision, recall, and F1-score are reported in Table V. In the second experiment, we used public health dataset for the training and Cleveland for the evaluation purpose. We achieved a remarkable accuracy of 99.34%. In the third experiment, we used the public health dataset for training purpose and Statlog for testing purpose and vice versa. We achieved an accuracy of 70.74%. In the fourth experiment, we used Statlog dataset for training purpose and public health dataset for evaluation purpose. We achieved an accuracy of 64.88%. From the results reported in Table V, we observed that the performance of the proposed system degrades when training on Statlog and testing on the public health dataset. We analyzed both datasets and found that the data has few issues. The number of instances of Statlog and public health datasets are either missing or very less as compared to another dataset. The amount of training data is small, and we can't use augmentation for the clinical data. The instances with fewer numbers are called the rare cases. We investigated both the datasets deeply and discovered that there are missing data in the training dataset (Statlog) for chest pain (cp), major vessels (ca), and thal while the testing dataset (public health dataset) has data for these features. There are four types of chest pain i.e, typical angina, atypical angina, non-anginal pain, asymptomatic. Typical angina is represented by 1, atypical angina by 2, non-anginal by 3, and asymptomatic by 4. From the dataset, we observed that there are a number of major vessels (0–3) colored by fluoroscopy. We analyzed that there are three types of thal i.e., normal, fixed defect, and reversible defect. Normal thal is denoted by 3, fixed defect is denoted by 6, and the reversible defect is denoted by 7. An over-constrained model underfits when there is a small amount of training data, whereas, an under-constrained model overfits the training data. Both of these cases result in poor prediction of the model. So, we conclude that the small amount of data for training in Statlog dataset is the main reason behind the poor prediction.

Table V. Performance results on cross datasets.

Training Dataset	Testing Dataset	Accuracy%	Precision%	Recall%	F1-Score%
Cleveland	Public Health Dataset	98.44	97.05	100	98.50
Public Health Dataset	Cleveland	99.34	98.80	100	99.40
Public Health Dataset	Statlog	70.74	86.60	56.00	68.02
Statlog	Public Health Dataset	64.88	59.93	92.25	7357

3.6. Performance comparison with other methods

To show the efficiency of our method for cardio disease prediction, we performed a comparative analysis of the proposed and existing state-of-the-art cardiovascular disease prediction methods, and results are shown in Table VI for Z-Alizadeh sani [42], public health [41], and Framingham [43] datasets, respectively. Our method yielded the best accuracy of 93.44%, 98.68%, and 99.86% on Z-Alizadeh sani, public health, and Framingham datasets, respectively. We have reported the results of comparative papers in this performance comparison experiment. First, we performed an experiment on Z-Alizadeh sani dataset

[42] and split the data into 80 – 20 for training and evaluation purposes. We used 80% data for training and the rest 20% for the testing. The results of the existing state-of-the-art methods [44–48] and our method are listed in Table VI. From Table VI, we can observe that [47] performed the worst and achieved an accuracy of 88.49%, abdar, [44] performed second best and yielded an accuracy of 93.08% while our method performed the best that yielded an accuracy of 93.44% on Z-Alizadeh sani dataset. The results reported in Table 6 reveals that the (RFE_XGboost) on eight features has better performance in terms of an accuracy and can be used reliably for the detection of cardiovascular disease.

Table VI. Performance comparison with other state-of-the-art methods on Z-Alizadeh Sani, Public health, and Framingham datasets.

Dataset	Authors	Accuracy%	Dataset	Authors	Accuracy%	Dataset	Authors	Accuracy%
Z-Alizadeh sani Dataset	Abdar, Moloud, et al. [44]	93.08	Public Health Dataset	Khan, Mohammad Ayoub et al. [49]	97.6	Framingham Health Dataset	Khan, Mohammad Ayoub et al. [49]	92.02
	Joloudari, Javad Hassannataj, et al. [45]	91.47		Khan, Mohammad Ayoub et al. [49]	84.6		Khan, Mohammad Ayoub et al. [49]	88.3
	Ghiasi, et al. [46]	92.41		Wu CS, et al. [50]	86		Sivaji, U., et al. [54]	88.7
	Khan et al. [47]	88.49		Ismail, A. et al. [51]	90.6		Al-Makhadmeh, et al. [55]	99.03
	Nasarian et al. [48]	92.35		Nurtas, Marat, et al. [52]	82		Nourmohammadi-Khiarak, Jalil, et al. [56]	94.03
	Nasarian et al. [48]	92.85		Raza, K, et al. [53]	88.88		Ali, Liaqat, et al. [57]	89
	In this study	**93.44**		**In this study**	**98.68**		**In this study**	**99.86**

Next, we performed an experiment on public health dataset [41] using 80 – 20 data for the training-testing. We compared the performance of our method with existing state-of-the-art methods as shown in Table VI. The detailed results of the proposed and existing state-of-the-art methods [49–53] in terms of an accuracy are reported in Table VI. From these results, we observed that [52] proposed a system for the prediction of heart disease and achieved an accuracy of 82%, which is 16.68% smaller than our method, [49] yielded second-best accuracy of 97.6% while our method achieved remarkable accuracy of 98.68%. The results reported in Table VI reveal that our experimental study is remarkably effective for the prediction of cardiovascular disease.

Finally, we used Framingham dataset [43] for the prediction of cardiovascular disease using 80% of the data for training the model and 20% of the data for testing. We compared the results of the proposed method against these contemporary methods [49, 54, 55, 56, 57] based on the accuracy as shown in Table VI. Experimental results on Framingham dataset [43] revealed that [49] achieved the worst performance with an accuracy of 88.3%, [55] achieved an accuracy of 99.3% while our method performs the best and yielded an accuracy of 99.86%. Experimental results show that the proposed system can effectively and reliably be used for the prediction of cardiovascular disease on multiple and diverse datasets.

4. Conclusion

This paper has presented an effective framework CNN-cardioAssistant to predict cardiovascular disease. Initially, we employed the RFE for selecting the most prominent features from the five diverse heart disease datasets i.e., Statlog, Cleveland, Public Health, Z-Alizadeh, and Framingham. Subsequently, we proposed a novel convolutional neural network for classification purpose and achieved an accuracy of 88.52%, 98.04%, and 99.58% on Z-Alizadeh, Public health, Framingham datasets. We evaluated the performance of the proposed method on cross datasets to check the generalizability of our method. In future, we aim to employ the same method for the prediction of other diseases. The RFE and other feature selection techniques with a combination of deep learning and machine learning classifiers can be applied to select more relevant features from the dataset to further enhance the performance of prediction systems.

Declarations

Acknowledgment

This work is supported by the Multimedia Signal Processing research lab at UET Taxila. Moreover, we would like to thank the UCI repository for providing the datasets.

Conflict of Interest

The authors declare that we don’t have any conflict of interest.

Ethics approval and consent to participate

Not applicable.

Availability of data and materials

The data is available on the UCI repositories. Each dataset is available on [39-43].

Funding

Funding is not available.

Author's Contributions

Farman Hassan contributed in the implementation, Auliya Ur Rahman Contributed in the manuscript write up, Ali Javed supervised the project, Ali Alhazmi helped in the proofreading, and Majed Alhazmi contributed in the manuscript writeup. All the authors contributed equally.

References

1. coronary artery disease. https://my.clevelandclinic.org/ health/diseases/16898-coronary-artery-disease.
2. Bonow R. O., Mann D. L., Zipes D. P., and Libby. Braunwald's P. (2012), heart disease: A textbook of Cardiovascular Medicine," Vol. 9, Saunders, New York.
3. Risk Factors for Coronary Artery Disease. https://www. healthline.com/health/coronary-artery disease/riskfactors.
4. Bache K., and Lichman M. (2013), "UCI machine learning repository," University of California, School of Information and Computer Science, Irvine, CA.
5. Durgesh K., and Bhambhu S. L., (2009), "Data classification using support vector machine," J. Theor. Appl. Inf. Technol., Vol. 12, pp. 1–7.
6. Tama B. A. and Rhee K.-H., (2019), "Tree-based classifier ensembles for early detection method of diabetes: an exploratory study," Artificial Intelligence Review, vol. 51, no. 3, article 9565, pp. 355–370.
7. Van Gestel T, Suykens JA, Baesens B, Viaene S, Vanthienen J, Dedene G, De Moor B, Vandewalle J. (2004), Benchmarking least squares support vector machine classifiers. Machine learning. Jan;54(1):5–32.
8. LeCun Y, Bengio Y, Hinton G. 2015, Deep learning. nature 521 (7553), 436–444. Google Scholar Google Scholar.
9. Domingos P, Pazzani M. (1997), On the optimality of the simple Bayesian classifier under zero-one loss. Machine learning.;29(2):103–30.
10. Sohn SY, Kim JW. (2012), Decision tree-based technology credit scoring for start-up firms: Korean case. Expert Systems with Applications.;39(4):4007–12.
11. Gao D, Zhang YX, Zhao YH. (2009) Random Forest algorithm for classification of multiwavelength data. Research in Astronomy and Astrophysics.;9(2):220.
12. Gou J, Ma H, Ou W, Zeng S, Rao Y, Yang H. A (2019) generalized mean distance-based k-nearest neighbor classifier. Expert Systems with Applications.; 115:356–72.
13. Ayon SI, Islam M. (2019), Diabetes Prediction: A Deep Learning Approach. International Journal of Information Engineering & Electronic Business.;11(2).
14. Manogaran G, Varatharajan R, Priyan MK. (2018) Hybrid recommendation system for heart disease diagnosis based on multiple kernel learning with adaptive neuro-fuzzy inference system. Multimedia tools and applications. 77(4):4379–99.
15. Hasan MK, Islam MM, Hashem MM. (2016) Mathematical model development to detect breast cancer using multigene genetic programming. In2016 5th International Conference on Informatics, Electronics and Vision (ICIEV) (pp. 574–579). IEEE.
16. Haque MR, Islam MM, Iqbal H, Reza MS, Hasan MK (2018). Performance evaluation of random forests and artificial neural networks for the classification of liver disorder. In2018 international conference on computer, communication, chemical, material and electronic engineering (IC4ME2) (pp. 1–5). IEEE.
17. Islam MM, Iqbal H, Haque MR, Hasan MK. (2017), Prediction of breast cancer using support vector machine and K-Nearest neighbors. In2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC) (pp. 226–229). IEEE.
18. Kadam K, Kamat PV, Malav AP. (2019) cardiovascular disease prediction using data mining techniques: A proposed framework using big data approach. InCoronary and Cardiothoracic Critical Care: Breakthroughs in Research and Practice (pp. 246–264). IGI Global.
19. Mathan K, Kumar PM, Panchatcharam P, Manogaran G, Varadharajan R. (2018) A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. Design automation for embedded systems.;22(3):225–42.
20. Shylaja S, Muralidharan R. (2018) Comparative analysis of various classification and clustering algorithms for heart disease prediction system. Biometrics Bioinf.;10(4):74-.
21. Singh P, Singh S., and Pandi-Jain, G. S. (2018), "Effective heart disease prediction system using data mining techniques," Int. J. Nanomed., Vol. 13, no. 121, pp. 47–70.
22. Ali F, El-Sappagh S, Islam SR, Kwak D, Ali A, Imran M, Kwak KS. (2020), A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion.; 63:208–22.
23. Sharma, Sumit, and Mahesh Parmar. (2020), "Heart diseases prediction using deep learning neural network model." International Journal of Innovative Technology and Exploring Engineering (IJITEE) 9.3.
24. Guan W, Gray A, Leyffer S. (2009), Mixed-integer support vector machine. InNIPS workshop on optimization for machine learning.
25. Shilaskar S, Ghatol A. (2013) Feature selection for medical diagnosis: Evaluation for cardiovascular diseases. Expert Systems with Applications.;40(10):4146–53.
26. Shao YE, Hou CD, Chiu CC. (2014) Hybrid intelligent modeling schemes for heart disease classification. Applied Soft Computing.; 14:47–52.
27. Inbarani HH. (2015) A novel neighborhood rough set-based classification approach for medical diagnosis. Procedia Computer Science.; 47:351–9.
28. Rajathi S, Radhamani G. (2016) Prediction and analysis of Rheumatic heart disease using kNN classification with ACO. In2016 International Conference on Data Mining and Advanced Computing (SAPIENCE) (pp. 68–73). IEEE.
29. Bashir S, Qamar U, Javed MY. (2014) An ensemble-based decision support framework for intelligent heart disease diagnosis. InInternational conference on information society (i-Society 2014) (pp. 259–264). IEEE.

30. Amin SU, Agarwal K, Beg R. (2013) Genetic neural network-based data mining in prediction of heart disease using risk factors. In2013 IEEE Conference on Information & Communication Technologies (pp. 1227–1231). IEEE. Information society (i-Society 2014) (pp. 259–264). IEEE.
31. Mehmood, Awais, et al. (2021) "Prediction of Heart Disease Using Deep Convolutional Neural Networks." Arabian Journal for Science and Engineering 46.4: 3409–3422.
32. Er O, Tanrikulu AC, Abakay A, Temurtas F. (2012) An approach based on probabilistic neural network for diagnosis of Mesothelioma's disease. Computers & Electrical Engineering.;38(1):75–81.
33. Er O, Yumusak N, Temurtas F. (2010) Chest diseases diagnosis using artificial neural networks. Expert Systems with Applications.;37(12):7648–55.
34. Liu X, Tosun D, Weiner MW, Schuff N, (2013) Alzheimer's Disease Neuroimaging Initiative. Locally linear embedding (LLE) for MRI based Alzheimer's disease classification. Neuroimage.; 83:148–57.
35. Nashif S, Raihan MR, Islam MR, Imam MH. (2018) Heart disease detection by using machine learning algorithms and a real-time cardiovascular health monitoring system. World Journal of Engineering and Technology. 12;6(4):854 – 73.
36. Singh M, Martins LM, Joanis P, Mago VK. (2016) Building a cardiovascular disease predictive model using structural equation model & fuzzy cognitive map. In2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) 2016 Jul 24 (pp. 1377–1382). IEEE.
37. Ghadge P, Girme V, Kokane K, Deshmukh P. (2015) Intelligent heart attack prediction system using big data. International Journal of Recent Research in Mathematics Computer Science and Information Technology.;2(2):73–7.
38. Guyon I, Weston J, Barnhill S, Vapnik V. (2002) Gene selection for cancer classification using support vector machines. Machine learning.;46(1):389–422.
39. Access on 7/14/2021 Available online at: Statlog Dataset. http://archive.ics.uci.edu/ml/datasets/statlog+(heart).
40. Access on 7/14/2021 Available online at: Cleveland Dataset.https://archive.ics.uci.edu/ml/datasets/heart+disease.
41. Access on 7/14/2021 Available online at: Public health Dataset. https://www.kaggle.com/ronitf/heart-disease-uci.
42. Access on 7/14/2021 Available online at: Z-aliZadeh Dataset. https://archive.ics.uci.edu/ml/datasets/Z-Alizadeh+Sani
43. Access on 7/14/2021 Available online at: Framingham Dataset. https://biolincc.nhlbi.nih.gov/studies/framcohort/.
44. Abdar M, Książek W, Acharya UR, Tan RS, Makarenkov V, Pławiak P. (2019) A new machine learning technique for an accurate diagnosis of coronary artery disease. Computer methods and programs in biomedicine. 1; 179:104992.
45. Joloudari JH, Hassannataj Joloudari E, Saadatfar H, Ghasemigol M, Razavi SM, Mosavi A, Nabipour N, Shamshirband S, Nadai L. (2020) Coronary artery disease diagnosis; ranking the significant features using a random trees model. International journal of environmental research and public health.;17(3):731.
46. Ghiasi MM, Zendehboudi S, Mohsenipour AA. (2020) Decision tree-based diagnosis of coronary artery disease: CART model. Computer methods and programs in biomedicine. 1; 192:105400.
47. Khan Y, Qamar U, Asad M, Zeb B. (2019) Applying feature selection and weight optimization techniques to enhance artificial neural network for heart disease diagnosis. InProceedings of SAI Intelligent Systems Conference 5 (pp. 340–351). Springer, Cham.
48. Nasarian E, Abdar M, Fahami MA, Alizadehsani R, Hussain S, Basiri ME, Zomorodi-Moghadam M, Zhou X, Pławiak P, Acharya UR, Tan RS. (2020) Association between work-related features and coronary artery disease: A heterogeneous hybrid feature selection integrated with balancing approach. Pattern Recognition Letters. 1; 133:33–40.
49. Khan MA. (2020) An IoT framework for heart disease prediction based on MDCNN classifier. IEEE Access. 17; 8:34717–27.
50. Wu CS, Badshah M, Bhagwat V. (2019) Heart Disease Prediction Using Data Mining Techniques. InProceedings of the 2019 2nd International Conference on Data Science and Information Technology 19 (pp. 7–11).
51. Ismail A, Abdlerazek S, El-Henawy IM. (2020) Big data analytics in heart diseases prediction. Journal of Theoretical and Applied Information Technology. 15;98(11).
52. Nurtas M, Baishemirov Z, Tsay V, Tastanov M, Zhanabekov Z. (2020) APPLYING NEURAL NETWORK FOR PREDICTING CARDIOVASCULAR DISEASE RISK. Известия НАН РК. Серия физико-математических наук. 12(4):28–34.
53. Raza K. (2019) Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule. InU-Healthcare Monitoring Systems 1 (pp. 179–196). Academic Press.
54. Sivaji U, Rao NK, Srivani C, Sree T, Singh M. (2021) A Hybrid Random Forest Linear Model approach to predict the heart disease. Annals of the Romanian Society for Cell Biology. 5;25(6):7810-4.
55. Al-Makhadmeh Z, Tolba A. (2019) Utilizing IoT wearable medical device for heart disease prediction using higher order Boltzmann model: A classification approach. Measurement. 1; 147:106815.
56. Nourmohammadi-Khiarak J, Feizi-Derakhshi MR, Behrouzi K, Mazaheri S, Zamani-Harghalani Y, Tayebi RM. (2020) New hybrid method for heart disease diagnosis utilizing optimization algorithm in feature selection. Health and Technology.;10(3):667–78.
57. Ali L, Rahman A, Khan A, Zhou M, Javeed A, Khan JA. (2019) An automated diagnostic system for heart disease prediction based on ${\chi^ {2}} $ statistical model and optimally configured deep neural network. IEEE Access. 13; 7:34938–45.
58. LeCun, Y.; Bengio, Y.; Hinton, G.: Deep learning. Nature 521, 436–444, 2015.
59. LeCun, Y.; Bengio, Y.: Convolutional networks for images, speech, and time series. Handbook Brain Theory Neural Netw. 3361, 1995, 1995.
60. Iandola, F.: Exploring the design space of deep convolutional neural networks at large scale, arXiv preprint arXiv:1612.06519 (2016).
61. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A.: Inception-v4, inception-resnet and the impact of residual connections on learning. In: Thirty-First AAAI Conference on Artificial Intelligence, (2017).

62. He, K.; Zhang, X.; Ren, S.; Sun, J.: Deep residual learning for image recognition. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, ed: IEEE (2016).

63. Krizhevsky, A.; Sutskever, I.; Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012).

64. Radford, A.; Metz, L.; Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434 (2015).

Figures

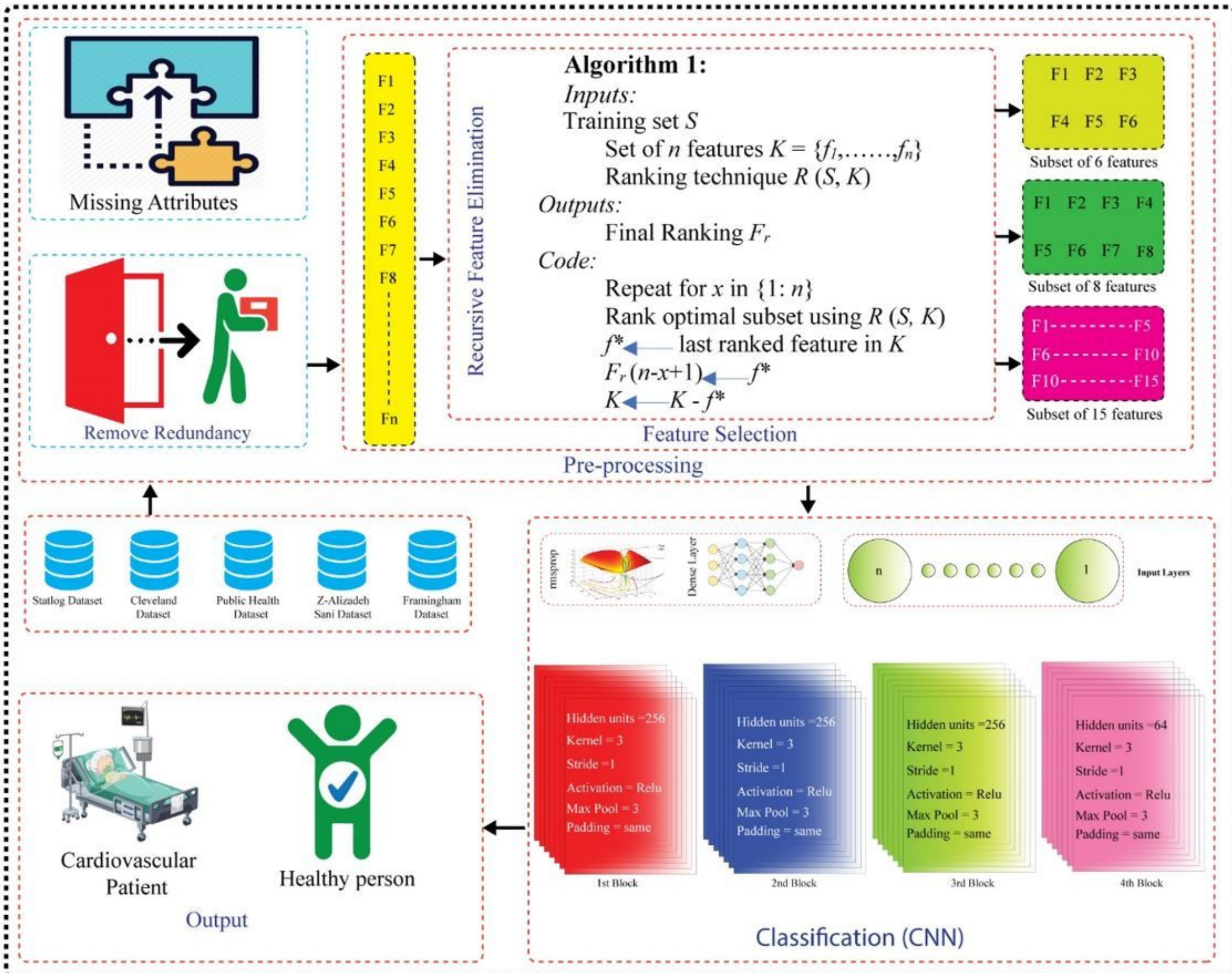

Figure 1

Proposed System.

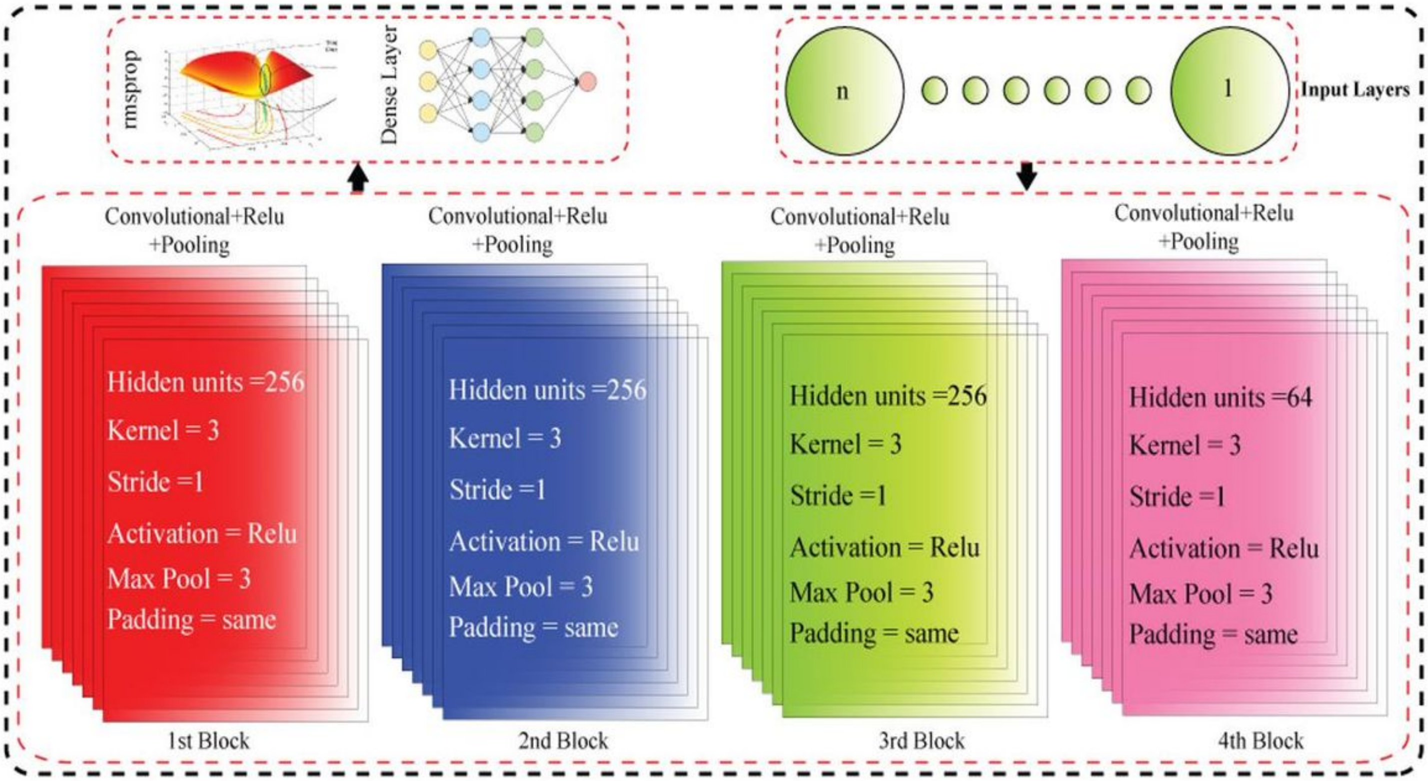

Figure 2

Proposed Convolutional Neural Network.