IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 2, 2020, accepted March 12, 2020, date of publication March 16, 2020, date of current version April 7, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2981159

Recursion Enhanced Random Forest With an Improved Linear Model (RERF-ILM) for Heart Disease Detection on the Internet of Medical Things Platform

CHUNYAN GUO [1], JIABING ZHANG [2], YANG LIU [1], YAYING XIE [1], ZHIQIANG HAN [1], AND JIANSHE YU [1]

[1]Anesthesiology Department, The Affiliated Hospital of Inner Mongolia Medical University, Hohhot 010050, China
[2]Graduate School of Medical School of Chinese PLA Hospital, Beijing 100853, China

Corresponding author: Jianshe Yu (yujianshe@nmgfy.com)

This work was supported by the Inner Mongolia Medical University through the ''One Million Yuan Project of Science and Technology'' under Grant YKD2018KJBW(LH)022.

ABSTRACT Nowadays, Heart disease is one of the crucial impacts of mortality in the country. In clinical data analysis, predicting cardiovascular disease is a primary challenge. Deep learning (DL) has been demonstrated to be effective in helping to determine and forecast a huge amount of data produced by the health industry. In this paper, the proposed Recursion enhanced random forest with an improved linear model (RFRF-ILM) to detect heart disease. This paper aims to find the key features of the prediction of cardiovascular diseases through the use of machine learning techniques. The prediction model is adding various combinations of features and various established methods of classification. it produces a better level of performance with precision through the heart disease prediction model. In this study, the factors leading to cardiovascular disease can be diagnosed. A comparison of important variables showed with the Internet of Medical Things (IoMT) platform, for data analysis. This indicates that coronary artery disease develops more often in older ages. Also important in this disease's outbreak is high blood pressure. For this purpose, measures must be taken to prevent this disease and Diabetes provides a further aspect that should be taken into consideration in the occurrence of coronary artery disease with 96.6 % accuracy,96.8% stability ratio and 96.7% F-measure ratio.

INDEX TERMS Heart disease detection, linear model, random forest, machine learning, diagnosis.

I. INTRODUCTION

Heart disease is a collection of diseases impacting the heart and veins of human beings. Cardiac disease symptoms vary depending on the specific type of cardiac disease [1]. Detecting and diagnosing the cardiovascular disease is an on-going job that can be achieved with enough experience and knowledge by a qualified professional [2]. There are many factors including age, diabetes, smoking, overweight, junk foods diet and so on. Several factors/parameters have been identified that cause heart disease or increase cardiac disease [3]. Most hospitals have management software for monitoring their clinical and/or patient data. It is popular now and Such systems produce enormous amounts of patient information. These data are seldom used for clinical decision-making support [4]. These data are valuable and information is kept largely unused in these data. It is an extremely difficult task to turn the accumulated clinical data into useful information that can make intelligent systems support decision-making for healthcare practitioners [5]. This factor led to research on the processing of medical pictures Due to the lack of experts and the number of cases incorrectly diagnosed, a rapid and efficient automated detection system was required. The main purpose is to classify the key features of the medical data using the classifier model and use the models for the early prediction of cardiac disease [6].

As shown in Figure.1, deep learning is a machine learning approach that utilizes multiple neural network layers and vast

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Wei.

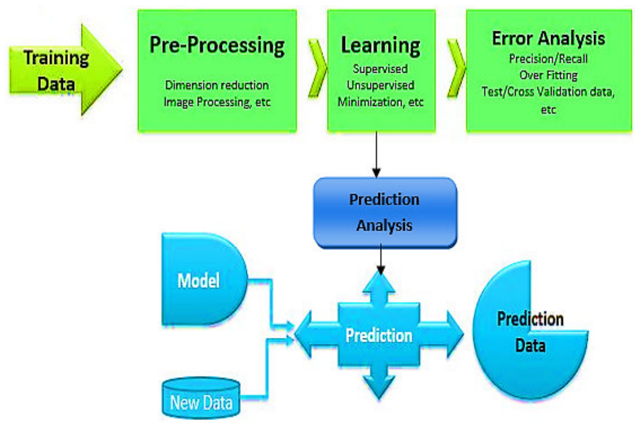


FIGURE 1. **Machine Learning model for image prediction.**

amounts of data to optimize a host of algorithms for a specific task [7]. The potential for therapeutic development, from discovery to prediction, to decision-making is high in machine learning. Within patient electronic healthcare records, ML can detect patterns of certain diseases and inform clinicians of all abnormalities [8]. Due to several contributory risk factors like diabetes, blood pressure, abnormal pulse rate, high cholesterol, and several other factors, it is critical to Machine Learning Model for image processing determines heart disease [9] Different techniques were employed to reduce the severity of cardiac disease among humans in the mining of data and neural networks [10]. Neural networks are widely considered to be the best way to predict diseases such as brain disease and heart disease. The generated results using the ANN which delivers good results in heart disease prediction [11]. Figure 1.1 shows the machine learning model in heart disease prediction. Neural network models incorporate not only later probabilities but expected values from several previous techniques. Neural network methods were implemented. The heart data set Cleveland with a neural network is used for all experiments to increase the performance of cardiac diseases [12].

Furthermore, The utilized Internet of Medical Things (IoMT) [13] is the array of healthcare devices and applications linked by online computer networks to healthcare IT systems. Medical devices that are equipped with Wi-Fi allow IoMT to communicate machine-to-machine [14], [15] approach. In IoMT systems, the data collected can be saved and analyzed via the cloud platforms [16], like the Amazon-WebServices [17]. IoMT is also referred to as IoT medical care [18]. Based on the survey the contribution has been listed as follows,

The main contribution of this paper,

- To evaluate the accuracy in the prediction of heart disease using Recursion Enhance Random Forest with an improved linear model (RFRF-ILM) method has been proposed.
- Designing an Artificial Neural Network with feature selection and backpropagation learning technique for classification of cardiovascular disease.
- Experimental results, the UCI Machine Learning Repository data set [19], [20] has been utilized for the training and testing for the performance evaluation.

The remainder of the paper represented as follows: Section 1 and Section 2 discussed the existing method for predicting heart disease and the theoretical method has been illustrated. In section 3 the Recursion Enhance Random Forest with an improved linear model (RFRF-ILM) method with classification modeling has been illustrated. In section 4 the experimental result has been discussed. Finally, section 5 concludes the research paper.

II. BACKGROUND SURVEY AND ITS SIGNIFICANCE IMPORTANCE

Health monitoring systems are essential components of smart health in the era of intelligent societies. In recent years, remote heart surveillance conditions for advanced machines have been developed to use advanced methods of electrocardiogram (ECG) to detect cardiac disorders. Current technologies are however affected by two major problems. i) failure to predict cardiac abnormalities to progress and (ii) failure to detect inter-patient variability. The author introduced a new prediction model for ECG signaling in two phases, which identifies significant abnormalities (red alarms) through the comparison of signals with a Global Classifier (GC). The proposed method has less predictive accuracy and has a special benefit to predictive analyzes by creating warning messages on the high risk of heart defects for medical care [21].

The growth of the economy and the mechanized way of life, and the disease incidence are also steadily increasing. It is predicted that some 23.7 million people will die from cardiovascular disease (CVD) within the next 10 years continuing the trend and lifecycle trend. In [20] Deep learning (DLTs) is therefore designed to evaluate stable CVDs to reduce RHI mal diagnosis. This means that the CVDs need to be analyzed. The goal of this paper is, to synthesize and identify CVD patients who entered the emergency section from January (2018 to December 2019) with molecular diagnostics (MD), and then with Deep Learning Techniques (DLTs).

The medical industries generate large quantities of data containing hidden information that is helpful in making successful decision-making. Such sophisticated techniques are used in data mining to produce adequate results and to make effective data decisions. The study uses the Neural Network to evaluate the heart disease's risk level to establish an effective heart disease prediction system (EHDPS) [21]. The system uses 15 clinical elements such as age, gender, blood pressure, cholesterol and obesity to predict this condition. The EHDPS measures the risk of cardiovascular disease in patients. It allows for important knowledge, for example, the relationships between heart disease-related medical factors and the patterns. The results showed that the diagnostic system developed can effectively predict the level of risk for heart disease.

Heart disease prediction is one field in which machine learning can be implemented. Several Optimization

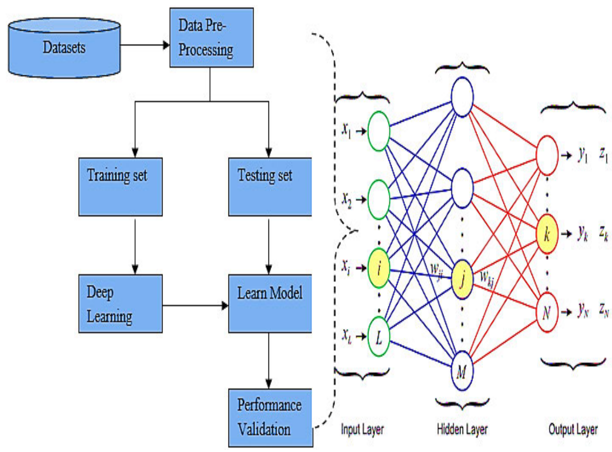


FIGURE 2. Deep Learning Technique for Prediction of heart disease.

algorithms have the advantage of being very flexible and adaptable to deal with complex, non-linear problems. The author introduced a fast-correlation-based selection method (FCBF) to filter redundant functions to enhance the quality of classification of heart disease. Then they classify based upon various classification modeling, like Support Vector Machine, K-Nearest Neighbor, Naïve Bayes, Random Forest, and Multilayer Perceptive Artificial Neural Network optimized with Particle Swarm Optimization (PSO) and Ant Colony Optimization approaches (ACO) [22]. The suggested mixed-method applies to cardiac disease datasets; the results show the effectiveness and robustness in processing different types of cardiac disease classification data for the proposed hybrid method.

III. RECURSION ENHANCE RANDOM FOREST WITH AN IMPROVED LINEAR MODEL (RFRF-ILM) METHOD

In this paper, the proposed excursion Enhance Random Forest with an improved linear model (RFRF-ILM) to predict heart diseases. Nearly all systems with coronary heart disease have parameters and input from complicated studies in the diagnoses of heart disorder. However, some research has been done to eliminate the dangerous effect of heart disease. In this work, information for several cardiac patients, classification algorithms, has been accumulated to predict the cardiac disorder of the person concerned. The satisfactory classifier by calculating the exactness of different classifications has been analyzed in this paper. The highest accuracy prediction has been implemented with ANN in the medical field. Furthermore, Accuracy, reliability, and repeatability are the three important elements of IoMT that must be always prioritized as shown in Figure.2.

As inferred from Figure.2. Dataset clustering is based on Decision Tree (DT) feature variables and criteria. To estimate its performance, the classifier is then applied to each data set. Based on their less error rate, the good performing models are identities based on the results. By selecting the decision tree cluster with a high error rate and removing its respective class-type features, the output is further optimized. To optimize the error on this dataset, the classifier performance reset is calculated. In Artificial Neural Networks, an input is received at the start of the input layer as $xx_{1....L}$ and passed on to a modified Model of the next layer input. The layers between the input and output are called hidden layers with the weight w_{ji} ....w_{kj}, which consist of several straight and non-linear transformations with the output function ($y_{1........N}$, $Z_{1........N}$). The several prepositions explain the structural flow which is explained as follows,

TABLE 1. Attributes information.

Attributes	*Description*	*Value*
Age	*Age in years*	*39*
Trestbps	*Resting blood pressure*	*32*
Cp	*Chest pain type*	*04*
Sex	*Sex (01=male; 02=female)*	*02*
Fbs	*Fasting blood sugar> 120 mg/dl 01=true ; 02=false*	*02*
Chol	*Serum cholesterol in mg/dl*	*144*
Ca	*Number of major vessels (0-3) colored by fluoroscopy*	*05*
Restecg	*Resting electrocardiographic results*	*02*
Slope	*The slope of peak exercise ST segment*	*11*
Exang	*Exercise-induced angina*	*07*
Thal	*3=normal; 6=fixed defect; 7=reversable defect*	*05*
Oldpeak	*ST depression induced by exercise relative to rest*	*06*
Num	*Diagnosis of heart disease*	*05*
thalach	*Maximum heart rate achieved*	*74*

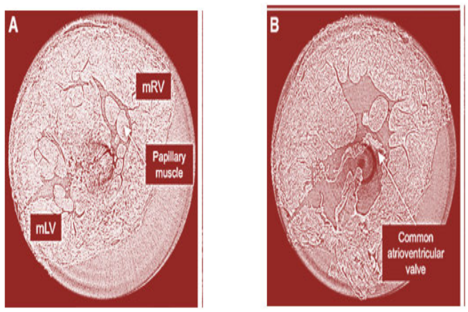
Preposition 1 (Decision Trees):

The main purpose of selection tree usage is for a practice version to be used for mastering the decision rules from the previous statistics to predict the magnificence or price of the target variables. A dataset is divided into smaller subsets and a related selection tree is progressively created. The tree with decision nodes and leaf nodes results from a decision tree. There may be two or additional branches of a decision node whereas the leaf node may be either a rank or a decision. The trees are the basis of high entropy inputs for training samples of data R. To eliminate the inappropriate samples on R, tree cutting is completed and the language model has been mathematically computed for resliced heart datasets.

$$Entropy = \sum_{i=1}^{n} q_{ji} log_2 q_{ji} \tag{1}$$

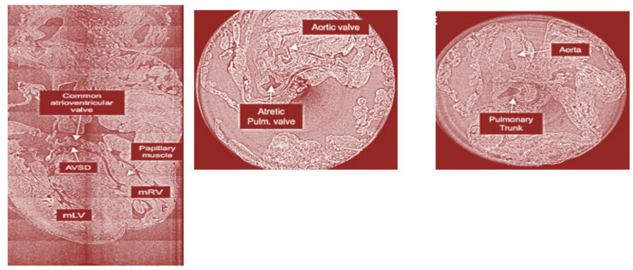
Preposition 2 (Language Model):

The linear solution type $f(y) = ny + a$ is resolved by the parameters has been used to identify the resilience heart structure as shown in Figure.3.(a) & (b) based on the



(a). Resliced Image.1. (b). Resliced Image.1.

FIGURE 3. **Computer Tomographic Resliced images.**



(a) papillary defect analysis (b) Tricuspid value (c) Ventricular

FIGURE 4. **Classifier for Septal identification a cardiovascular disease.**

following y_j, x_j input features with the data input vector y_j of data R.

$$n = \frac{\left(\sum_j y_j x_j\right) - m\overline{y_j x_j}}{\left(\sum_j y_j^2\right) - m\overline{y_j^2}} \tag{2}$$

As shown in equation (2) $a = \bar{x} - n\bar{y}$ where $\bar{y}, \bar{x}$ are the means. Further disorder prediction has been analyzed using the Vector machine approach.

Preposition 3 (Support Vector Machine (SVM)):

The SVM approach in disorder prediction has more correct and less error. Heart Disease Diagnosis SVM plays better with the greatest accuracy. Support Vector Support (SVM) is a supervised system that examines algorithms that can be used in situations demanding each class or each regression. Let the samples of training with data test $Data = \{x_j, y_j\}$; $j = 1, 2, \ldots m$, where $y_j \in G^m$ is the vector and $x_j \in G^m$ is the destination component. Linear SVM identifying the best way of hyperplane form $f\ (y) = s^T y + a$, where a is offset and s is a dimensional coefficient vector. This is accomplished by resolving the following problem optimization. Further, the decision making has been proposed through the random forest model as described below.

$$\begin{aligned} &Min_{s,a,\xi_j}\ \frac{1}{2}s^2 + H\sum_{j=1}^{m}\xi_j \\ &s.t.\ x_j\left(s^T y_j + a\right) \geq 1 - \xi_j, \quad \xi_j \geq 0,\ \forall j \in (1, 2, ..i) \end{aligned} \tag{3}$$

Preposition 4 (Random Forest):

As shown in Figure4 (a,b,c), this classifier creates and incorporates many decision-making trees to achieve better results. For tree learning, bagging or bootstrap aggregation mainly applies. The hidden samples y' is generated by averaging the predictions $\sum_{c=1}^{C} fc(y')$ from each tree on y' for the papillary, tricuspid and ventricular analysis. For a given data, $Y = \{y_1, y_2, y_3, \ldots y_m\}$ with responses $X = \{x_1, x_2, x_3, \ldots x_m\}$ which repeats the bagging form a = 1 to C.

$$i = \frac{1}{C}\sum_{c=1}^{C} fc(y') \tag{4}$$

The prediction uncertainty on this tree is expressed by its standard deviation,

$$\rho = \sqrt{\frac{\sum_{c=1}^{C}\left(fc\left(y'\right) - \hat{f}\right)^2}{C - 1}} \tag{5}$$

Preposition 5 (Naive Bayes):

Naive Bayes's model of learning employs Bayes rules utilizing separate features as analyzed by the naïve Bayes model. Each R instance is assigned the highest probability class subsequently. The model is trained with prior likelihood $Q\left(Y_f\right) = Priority \in (0:1)$. The model has a Gaussian function.

$$\begin{aligned} &Q\left(Y_{f1}, Y_{f2}, \ldots . Y_{fm}\,|h\right) \\ &\quad = \prod_{j=1}^{m} Q(Y_{fj}\,|h) \end{aligned} \tag{6}$$

$$\begin{aligned} &Q\left(Y_f\,|h_j\right) \\ &\quad = \frac{Q\left(h_j\,|Y_f\right)Q(Y_f)}{Q(h_j)} \quad h \in \{benign, malignant\} \end{aligned} \tag{7}$$

Finally, the test data is categorized according to association probability with neural networks for mathematical computation,

$$h_{ma} = argmaxQ(h_l)\prod_{j=1}^{m} Q(Y_{fj}\,|h_l)\ , \quad for\ l = 1, 2 \tag{8}$$

Preposition 6 (Neural Networks):

Let's consider input y_j; hidden layers and output x_j are the neuron components. The results are obtained by the activation function such as sigmoid and a constant bias an as shown in Figure.5.

$$f\left(a + \sum_{j=1}^{m} y_j v_j\right) \tag{9}$$

Preposition 7 (K-Nearest Neighbour):

As shown in the Algorithm.1. Here cardiovascular disease prediction for the infected human fetal has been taken for analysis from UCI datasets. Figure.6. extracts the data from Euclidean distance function samples $r\left(y_j, y_i\right)$ and most of the K-NN based on neural representations for the extraction of basal,mid-high,mid-low and apical layers.

$$r\left(y_j, y_i\right) = \sqrt{(y_{j,1} - y_{i,1})^2 + \ldots . + (y_{j,n} - y_{i,n})^2} \tag{10}$$

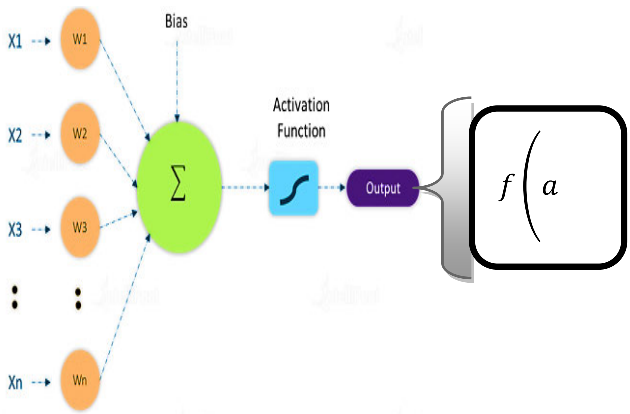


FIGURE 5. The activation function of the neural network representation.

Algorithm 1 Partition of Decision Tree

Input: Features Dataset R with the target group
For every sample do
For ∀ features do
Validate the algorithm of the decision tree
End for
Determine the feature space $f_1, f_2, \ldots f_y$, of the dataset
End for
Identify the leaf nodes total number $k_1, k_2, k_3, \ldots k_m$ with its equations
Divide the datasets R into $r_1, r_2, r_3, \ldots r_m$ basis of the leaf nodes equation.
Output: datasets Partition $r_1, r_2, r_3, \ldots r_m$

Apply the error rate-based hybrid approach to the following equation expressed in the algorithm.2.

As inferred from the algorithm.2,

$$\sum_0^m F(m) = r + n_1 y_1 + n_2 y_2 + \ldots n_m y_m \quad (11)$$
$$\sum_0^m F(0) = Gain + \sum_0^m s_j y_j \quad (12)$$

The proposed RFRF-ILM classification modeling achieves high accuracy in predicting heart disease. The classification modeling such as K-NN, Naive Bayes, ANN, Random forest, Decision tree, language model, etc. Further In this research paper, The proposed model is associated with deep learning and its basis of ANN, the learning can be unsupervised, supervised and semi-supervised. Here The pipeline method is used to classify a large amount of data in the cloud IoT system to achieve the performance and accuracy of data on the IoMT platform with the advanced data analysis approach for RFRF-ILM.

- Input data: A wide assortment of data can be utilized as a contribution to Deep learning purposes. This image data could come from an assortment of sources, for example, venture frameworks, centralized computer databases or IoT edge devices, and might be organized or unstructured in nature. Exceptionally high volumes of data are frequently fed into machine learners since more data often yields more bits of knowledge.
- Learning: Typically, the DL utilized for business reasons for existing is either unsupervised or supervised in nature. Inside these classes, however, there a wide range of type of algorithms and ML schedules, which can be utilized to achieve various objectives for analyzing the heart image.
- Output data: DL can be utilized to convey results that are either prescient or prescriptive. The outcomes can likewise convey yields that characterize heart image data or feature regions for investigation. This yield information may be put away for investigation, conveyed as reports or fed as a contribution to other undertaking applications or frameworks.

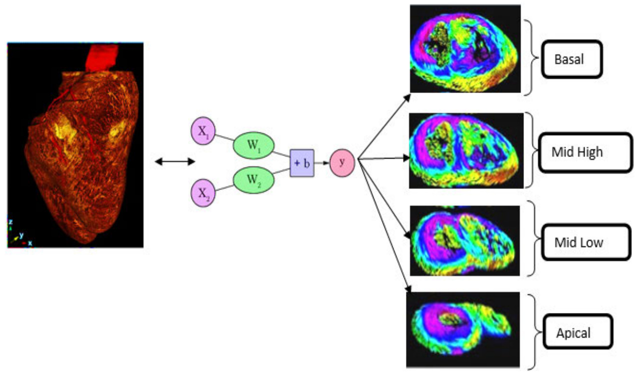


FIGURE 6. Fetal cardiovascular analysis using K-nearest neighbor with a neural network representation.

Algorithm 2 To Find the Least Rate of Error, Apply ML

Input: Datasets with partition
For ∀ employ the rules do
On the dataset $G(r_1, r_2, r_3, \ldots r_m)$
End for
dataset classify basis of the rules $H(G(r_1, G(r_2), \ldots G(r_n))$
Output: datasets classified with rules $H(G(r_1), G(r_2) \ldots G(r_m))$
For all datasets classified $H(G(r_1, G(r_2), \ldots G(r_n))$
For∀ determine min error rate from the input do
min $(H(G(r_1, G(r_2), \ldots G(r_n)))$
end for
Determine max(min) error rate from the classifier.
Output: features that have classifier $F(r_1, r_2, r_3, \ldots r_m)$

Indecision tree classifier for the IoMT environment, missing data are an important issue in a large amount of data clustering process. Let's consider the missing data value $G(q)$ the polynomial function can be expressed as the following equation (13) is,

$$G(a) = p_1 \frac{(a-a_2)(a-a_3)\ldots(a-a_m)}{(a_1-a_2)(a_1-a_3)\ldots(a_1-a_m)} + p_2 \frac{(a-a_1)(a-a_2)\ldots(a-a_m)}{(a_2-a_1)(a_2-a_3)\ldots(a_1-a_m)} + \ldots$$

$$+p_n \frac{(a-a_1)(a-a_2)\ldots(a-a_{n-1})}{(a_n-a_n)(a_n-a_2)\ldots(a_n-a_{n-1})} = \sum\nolimits_{j=0}^{m} p_j \prod\nolimits_{i=0,i\neq 0}^{m} \frac{a-a_i}{a_j-a_i} \quad (13)$$

In data analysis in IoMT resulting in the distance between the two different variable values. The data transform approach can be evaluated as the following equation (14) is,

$$a^* = \frac{a - min}{max - min} \quad (14)$$

As shown in the equation (2) where q is the data untransformed value and a^* denoted the normalized data value.

The evaluation of the mean and variance of data stationary tests is utilized. To be particular, the time series $\{A_d, d \in D\}$, is a consecutive value A_d the function of the autocorrelation coefficient of Q_d can be stated as the following equation (15(a),15(b)) is,

$$\delta(d, l) = H[(A_d - \mu_d)(A_l - \mu_l)] \quad (15(a))$$

$$\beta(d, l) = \frac{conv(A_d, A_l)}{\rho_d \rho_l} \quad (15(b))$$

The optimal data segmentation value is i and data segmentation focus is s, resolving (i, l) and the data label p_i for training sample as the following equation (16) is,

$$\min_{i,l}[\min_{b_1} \sum\nolimits_{q_i \in T_1(i,l)} (p_i - b_1)^2 \min_{b_2} + \sum\nolimits_{q_i \in T_2(i,l)} (p_i - b_2)^2 \quad (16)$$

As shown in the equation (16) the label function split into two inputs data as T_1, T_2 and estimate the mean value as the following equation (17) is,

$$T_1(i, l) = \left\{a | a^{(i)} \leq l\right\}, \quad T_2(i, l) = \left\{a | a^{(i)} > l\right\} \quad (17)$$

$$\hat{b}_n = \frac{1}{M_n} \sum\nolimits_{q_i \in T_m(i,l)} p_i, \quad a \in T_m, \; n = 1, 2. \quad (18)$$

As shown in the equation (18) where $\hat{b}_n$ is the mean value and T_m is the data region of p_j and a_j. The decision

To make a decision in the dataset the following equation (19) can be expressed by,

$$f(a) = \sum\nolimits_{n=1}^{N} \hat{b}_n K (a \in T_m), \quad (19)$$

The equation (20) where K is the indication function. Then the training samples of the mean value as the data predictive value represented as the following equation (20) is,

$$\hat{f} = \frac{1}{C} \sum\nolimits_{c=1}^{C} f_c(a) \quad (20)$$

In data analysis the observation training sample heart data matrix as,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1y} \\ a_{21} & a_{22} & \cdots & a_{2y} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{my} \end{bmatrix} \quad (21)$$

Algorithm 3 Decision Tree Classifier in RFRF-ILM

import input datasetG (q)
 # create a matrix including the data

$$\sum_{j=0}^{m} p_j \prod_{i=0,i\neq 0}^{m} \frac{q - q_i}{q_j - q_i}$$

 # defining the data frame from the matrix
$\{Q_d, d \in D\}$ respectively {data label, weight}
 # defining predictors

$$q | q^{(i)} \leq l$$

 # defining the target variable and mapping it to 1 for dog and 0 for cat
for q = 0, then

$$\hat{f} = \frac{1}{C} \sum_{c=1}^{C} f_c(q) = 1$$

else
q = 1 then

$$\hat{f} = \frac{1}{C} \sum_{c=1}^{C} f_c(q) = 0$$

 # instantiating the model
tree = Decision Tree Classifier ()

The standardizing the data in cloud clustering is evaluated as,

$$a_{ji}^* = \frac{a_{ji} - \bar{a}_i}{\sqrt{Var(a_i)}} \quad (i = 1, 2, \ldots, m; \; i = 1, 2, \ldots, y), \quad (22)$$

where,

$$\bar{a}_i = \frac{1}{m} \sum\nolimits_{j=1}^{m} a_{ji}$$

$$Var(a_i) = \frac{1}{m-1} \sum\nolimits_{j=1}^{m} (a_{ji} - \bar{a}_i)^2 \quad (i = 1, 2, \ldots y) \quad (23)$$

The training sample data correlation coefficient matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1y} \\ r_{21} & r_{22} & \cdots & r_{2y} \\ \vdots & \vdots & \ddots & \vdots \\ r_{y1} & r_{y2} & \cdots & r_{yy} \end{bmatrix} \quad (24)$$

The standardized data correlation coefficient of the original data is,

$$r_{ji} = \sum\nolimits_{k=1}^{m} (a_{kj} - \bar{a}_j)(a_{ki} - \bar{a}_i) / \sqrt{\sum\nolimits_{k=1}^{m} (a_{kj} - \bar{a}_j)^2 \sum\nolimits_{k=1}^{m} (a_{ki} - \bar{a}_i)^2} \quad (25)$$

The criterion of determining data components in the correlation coefficient matrix as the following equation (26) is,

$$\frac{\sum_{j=1}^{m} \gamma_j}{\sum_{j=1}^{y} \gamma_j} \geq \varphi \quad (26)$$

The evaluation of the data components as follows the equation (27) is,

$$X_j = \omega_{1j}A_1 + \omega_{2j}A_2 + \ldots + \omega_{yj}A_y, \quad j = 1, 2, \ldots n \tag{27}$$

Using this definition and theories IoMT environment problems in practical cases. The improvement of future IoMT database platform as shown in Figure.7 is permit to observe, react and learn in dynamic data management using deep learning assisted RFRF-ILM.

Here the Decision trees are one of the most mainstream calculations utilized in machine learning, for the most part for classification and additionally for data regression issues. When arranging a variable from training a heart image dataset, the possibility of the decision tree is to isolate the information into smaller datasets dependent on a specific component value unless the objective factors all be subjected to one class. Properties of the target function basis of the ''branches'' of the decision tree, the estimations of the target function are recorded in the ''leaves'', and the rest of the nodes contain features for which the cases vary.

The proposed RFRF-ILM method describes the three performance measuring initiatives that are used to better understand the behavior of different combinations in the selection of features to determine the important features of heart disease. DL technology concentrates on the good performing model compared to the other models. In heart disease prediction, the implement the RFRF-ILM which has high accuracy and low classic error. The quality of each class is individually assessed and all the outputs for further evaluation are properly recorded. Internet of Medical Things (IoMT) market contains smart devices like medical/vital monitors wearables, and the related Real-Time locations, telehealth, and other products specifically for use in healthcare on the body, at home and in societies. Figure 7 shows the IoMT flowchart used in this research paper.

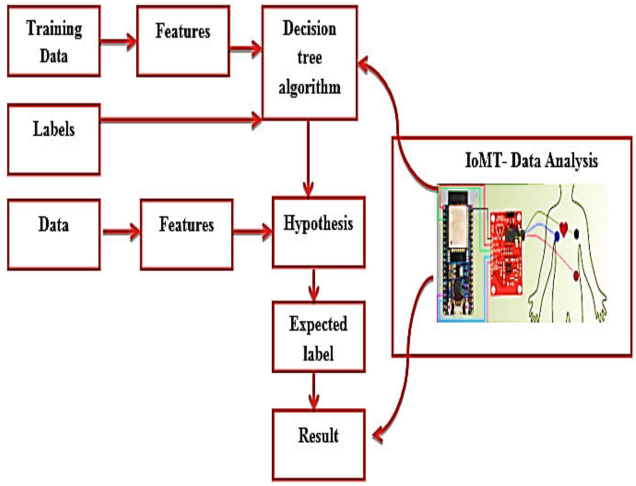


FIGURE 7. IoMT platform for training data analysis.

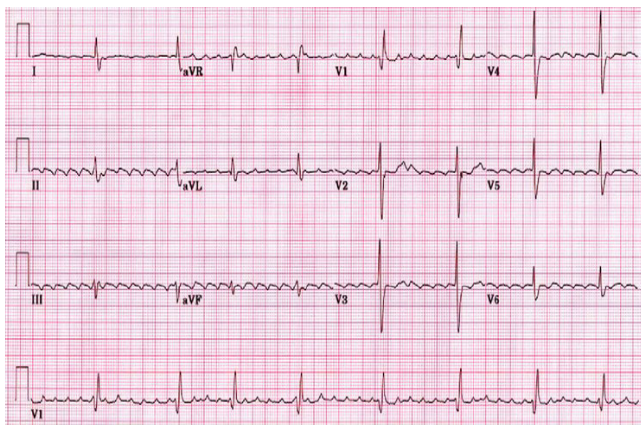


FIGURE 8. Cardiovascular disease dataset rhythm for atrial flutter [26].

IV. RESULTS AND DISCUSSION

Dataset clustering is based on Decision Tree (DT) feature variables and criteria. To estimate its Sensitivity and Specificity [24], [25], the classifier is then applied to each data set with the rhythm classification graphical structure as shown in the Figure.8 for arterial flutter has been analyzed whereas the dataset has been taken from administrative datasets of rom the collection of Dr K.C. Wu and [26]. Based on their less error rate, Further, the good performing models are identities based on the above results.

Sensitivity gives information on the percentage of correctly classified patients, while specificity provides data on properly classified healthy subjects. These assessment metrics are formulated as follows:

$$\textbf{\textit{Metric learing ratio}} = \frac{\boldsymbol{TP}}{\boldsymbol{TP+FN}} \tag{28}$$

$$\textbf{\textit{Outspace learning ratio}} = \frac{\boldsymbol{TN}}{\boldsymbol{TN+FP}} \tag{29}$$

where TP- True Positive, FP- False Positive, FN-False Negative, TN-True Negative

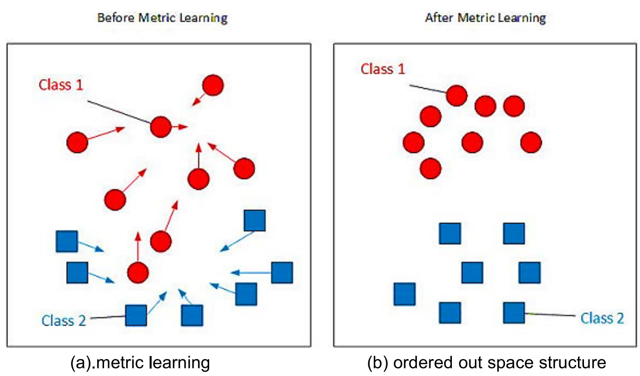


FIGURE 9. Metric Vs ordered out space analysis.

Figure 9(a) shows the metric learning of the proposed RERF-ILM. The proposed method has high sensitivity when compared to other existing methods such as GC, DLT, EHDPS, FCBF. Figure 9(b) shows the out space of the proposed RERF-ILM. The proposed method has high specificity when compared to other existing methods such as GC, DLT, EHDPS, FCBF.

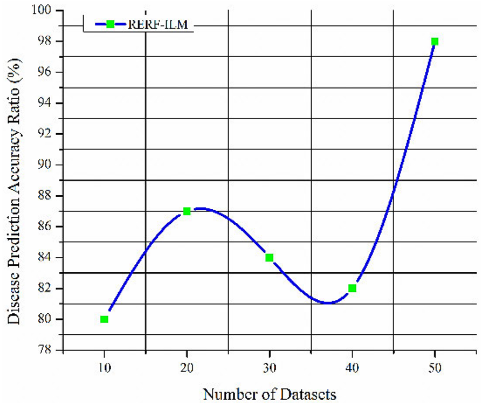


FIGURE 10. Cardiovascular disease analysis.

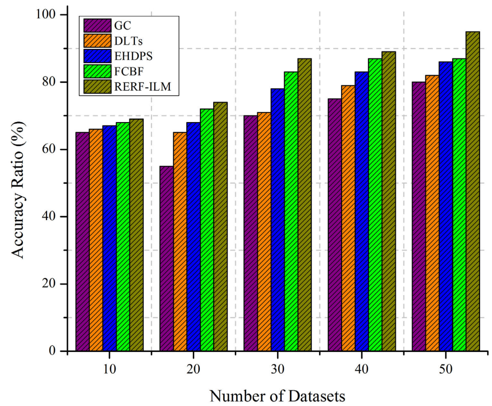


FIGURE 11. Accuracy analysis.

TABLE 2. Prediction ratio.

Number of Datasets	GC	DLTs	EHDPS	FCBF	RERF-ILM
10	65.1	66.7	67.4	68.8	69.5
20	55.6	65.5	68.3	72.8	74.9
30	70.8	71.1	78.9	83.2	87.4
40	75.3	79.3	83.5	87.3	89.8
50	80.9	82.7	86.2	87.9	95.2

TABLE 3. Validation stability ratio.

Number of Datasets	GC	DLTs	EHDPS	FCBF	RERF-ILM
10	65.5	67.6	68.4	69.7	70.1
20	75.6	77.9	83.4	88.6	92.1
30	70.2	74.6	77.8	79.8	82.3
40	75.6	78.9	83.4	86.7	90.6
50	80.9	84.6	87.9	90.7	98.8

As inferred from figure.9.(a,b) This classifier creates and incorporates many decision-making trees to achieve a better result in metric learning of heart disease prediction with ordered out space classes as shown in Figure.9. For tree learning, bagging or bootstrap aggregation mainly applies. It has also seen the use of decisional trees to predict the accuracy of heart disease-related events. Different methods for the prediction of heart disease were used for inference using machine learning methods. Several readings have been produced in this work for the development of a prediction model using not only different techniques but also two or more approaches. Table 2 shows the prediction ratio of the proposed RERF-ILM method.

The proposed RFRF-ILM method is utilized merging the features of the linear model and random forest. RFRF-ILM achieves high accuracy in the prediction of heart disease. The support vector machine is utilized to enhance the performance of the algorithm. The results of the suggested test case strategy showed better results than the results of other approaches for Cleveland heart disease as shown in Table 2. The prediction accuracy of the proposed method is presented in Figure 10.

Results from the Effective Heart Disease Prediction classification using the RERF-ILM method show that the accuracy of heart disease is more precise than with other methods (GC, DLT, EHDPS, FCBF). Figure 11 shows the accuracy of the proposed RERF-ILM method.

The compared classification results of our approach with existing GC, DLT, EHDPS, and FCBF methods used for data sets to determine the performance of the proposed approach with more accuracy. For a better comparison, the use of specific records that contain findings for all the literature methods. To avoid any distortion of the trial data set, the existing test data set to ensure that all samples are involved in the test process. Table 3 results indicate that the proposed method shows similar results in the prediction of cardiovascular disease and incorporates other approaches into the heart disease classification with high stability.

This paper aims to find the key features of the prediction of cardiovascular diseases through the use of machine learning techniques. The prediction model is adding various combinations of features and various established methods of classification. This model has a better level of performance with precision through the heart disease prediction model. Figure 12 demonstrates the performance ratio of the proposed method. Naive Bayes's model of learning employs Bayes rules utilizing separate features as analyzed by the naïve Bayes model. Each R instance is assigned the highest probability class subsequently with high performance.

Based on their less error rate, the good performing models are identities based on the above results. By selecting the decision tree cluster with a high error rate and removing its corresponding class-type features, the output is further optimized. To optimize the error on this dataset, the performance of the classifier reset is determined. Figure 13(a) demonstrates the error rate classification of the proposed RERF-ILM method and Figure 13(b) demonstrates the overall error rate of the proposed RERF-ILM method.

From the experimental results, this indicates that coronary artery disease develops more often in older ages and Also

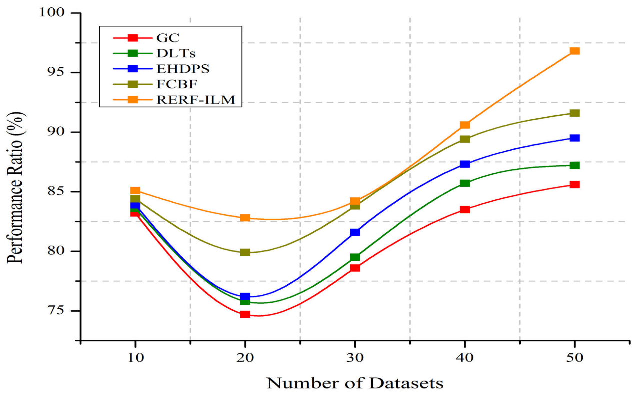
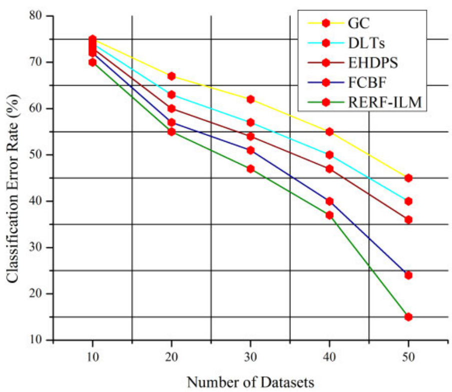
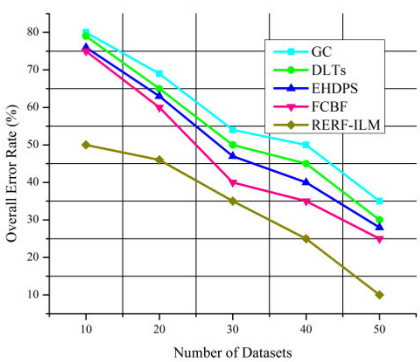


FIGURE 12. Performance Ratio.



(a) Classification Error Rate



(b) Overall Error Rate

FIGURE 13. Error rate analysis.

important in this disease's outbreak is high blood pressure provides a further aspect that should be taken into consideration in the occurrence of coronary artery disease.

V. CONCLUSION

The identification of raw health information processing will help to save lives and to detect abnormalities in cardiac conditions in the long term. In this study, machine-learning techniques have been utilized to manage raw data and to offer a better and new understanding of heart disease. The prediction of heart disease is very difficult and critical in the field of medicine. However, if the disease is a detective at an early stage and the prevention assessment is taken as quickly as possible, the mortality rate can be drastically controlled. The proposed RFRF-ILM method is utilized merging the features of the linear model and random forest. RFRF-ILM achieves high accuracy in heart disease prediction. The results achieved by the proposed method to test cases showed the best results than the results of other approaches. The support vector machine is utilized to enhance the performance of the algorithm. The algorithm proposed saves diagnostic costs and time and enhances the accuracy of the treatment process.

ACKNOWLEDGMENT

The Soothing Effect on Valve Replacement Patient on the Pulmonary Arterial Hypertension Via Transesophageal Echocardiography (TEE).

REFERENCES

[1] M. Kavousi, S. Elias-Smale, J. H. W. Rutten, M. J. G. Leening, R. Vliegenthart, G. C. Verwoert, G. P. Krestin, M. Oudkerk, M. P. M. de Maat, F. W. G. Leebeek, F. U. S. Mattace-Raso, J. Lindemans, A. Hofman, E. W. Steyerberg, A. van der Lugt, A. H. van den Meiracker, and J. C. M. Witteman, "Evaluation of newer risk markers for coronary heart disease risk classification: A cohort study," *Ann. Internal Med.*, vol. 156, no. 6, pp. 438–444, Mar. 2012.

[2] H. Tada, O. Melander, J. Z. Louie, J. J. Catanese, C. M. Rowland, J. J. Devlin, S. Kathiresan, and D. Shiffman, "Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history," *Eur. Heart J.*, vol. 37, no. 6, pp. 561–567, Feb. 2016.

[3] B. Narasimhan and A. Malathi, "Artificial lampyridae classifier (ALC) for coronary artery heart disease prediction in diabetes patients," *Int. J. Adv. Res., Ideas, Innov. Technol.*, vol. 5, no. 2, pp. 683–689, 2019.

[4] B. Narasimhan and A. Malathi, "Improved fuzzy artificial neural network (IFANN) classifier for coronary artery heart disease prediction in diabetes patients," *Indian J. Appl. Res.*, vol. 9, no. 4, pp. 1–4, 2019.

[5] V. Nambi, L. Chambless, M. He, A. R. Folsom, T. Mosley, E. Boerwinkle, and C. M. Ballantyne, "Common carotid artery intima-media thickness is as good as carotid intima-media thickness of all carotid artery segments in improving prediction of coronary heart disease risk in the atherosclerosis risk in communities (ARIC) study," *Eur. Heart J.*, vol. 33, no. 2, pp. 183–190, Jan. 2012.

[6] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, and A. Lamgunde, "Heart disease prediction using data mining techniques," in *Proc. Int. Conf. Intell. Comput. Control (ICICC)*, Jun. 2017, pp. 7–11.

[7] P. K. Anooj, "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 24, no. 1, pp. 27–40, 2012.

[8] J. L. Zulueta, V. L. Vida, E. Perisinotto, D. Pittarello, and G. Stellin, "The role of intraoperative regional oxygen saturation using near infrared spectroscopy in the prediction of low output syndrome after pediatric heart surgery," *J. Cardiac Surg.*, vol. 28, no. 4, pp. 446–452, Jul. 2013.

[9] P. S. de Vries, M. Kavousi, S. Ligthart, A. G. Uitterlinden, A. Hofman, O. H. Franco, and A. Dehghan, "Incremental predictive value of 152 single nucleotide polymorphisms in the 10-year risk prediction of incident coronary heart disease: The rotterdam study," *Int. J. Epidemiol.*, vol. 44, no. 2, pp. 682–688, Apr. 2015.

[10] S. Baskar, V. R. S. Dhulipala, P. M. Shakeel, K. P. Sridhar, and R. Kumar, "Hybrid fuzzy based spearman rank correlation for cranial nerve palsy detection in MIoT environment," *Health Technol.*, vol. 10, pp. 259–270, 2020, doi: 10.1007/s12553-019-00294-8.

[11] N. Gupta, N. Ahuja, S. Malhotra, A. Bala, and G. Kaur, "Intelligent heart disease prediction in cloud environment through ensembling," *Expert Syst.*, vol. 34, no. 3, Jun. 2017, Art. no. e12207.

[12] M. S. Sabatine, "Using aptamer-based technology to probe the plasma proteome for cardiovascular disease prediction," *JAMA*, vol. 315, no. 23, pp. 2525–2526, Jun. 2016.

[13] A. Wilson and G. Wilson, "Heart disease prediction using data mining techniques," *Int. J. Comput. Sci. Trends Technol.*, vol. 2, no. 1, pp. 84–88, 2014.

[14] A. K. Pandey, P. Pandey, and K. L. Jaiswal, "A novel frequent features prediction model for heart disease diagnosis," *Int. J. Softw. Hardw. Res. Eng.*, vol. 1, no. 1, pp. 38–43, 2013.
[15] G. C. M. Siontis, I. Tzoulaki, K. C. Siontis, and J. P. A. Ioannidis, "Comparisons of established risk prediction models for cardiovascular disease: Systematic review," *BMJ*, vol. 344, p. e3318, May 2012.
[16] R. Poplin, A. V. Varadarajan, K. Blumer, Y. Liu, M. V. McConnell, G. S. Corrado, and D. R. Webster, "Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning," *Nature Biomed. Eng.*, vol. 2, no. 3, p. 158, 2018.
[17] J. Wang, H. Ding, F. A. Bidgoli, B. Zhou, C. Iribarren, S. Molloi, and P. Baldi, "Detecting cardiovascular disease from mammograms with deep learning," *IEEE Trans. Med. Imag.*, vol. 36, no. 5, pp. 1172–1181, May 2017.
[18] P. Gomathi, S. Baskar, and P. M. Shakeel, "Identifying brain abnormalities from electroencephalogram using evolutionary gravitational neocognitron neural network," *Multimedia Tools Appl.*, vol. 2019, pp. 1–20, Feb. 2019, doi: 10.1007/s11042-019-7301-5.
[19] J. Chen, A. Valehi, and A. Razi, "Smart heart monitoring: Early prediction of heart problems through predictive analysis of ECG signals," *IEEE Access*, vol. 7, pp. 120831–120839, 2019.
[20] *Heart Disease Data Set*. [Online]. Available: https://archive.ics.uci.edu/ml/datasets/Heart+Disease
[21] A. Junejo, Y. Shen, A. A. Laghari, X. Zhang, and H. Luo, "Molecular diagnostic and using deep learning techniques for predict functional recovery of patients treated with cardiovascular disease," *IEEE Access*, vol. 7, pp. 120315–120325, 2019.
[22] P. Singh, S. Singh, and G. S. Pandi-Jain, "Effective heart disease prediction system using data mining techniques," *Int. J. Nanomed.*, vol. 13, p. 121, Mar. 2018.
[23] Y. Khourdifi and M. Bahaj, "Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization," *Int. J. Intell. Eng. Syst.*, vol. 12, no. 1, pp. 242–252, 2019.
[24] A. E. Di, P. Gao, H. Khan, A. S. Butterworth, D. Wormser, S. Kaptoge, and P. M. Ridker, "Glycated hemoglobin measurement and prediction of cardiovascular disease," *Jama*, vol. 311, no. 12, pp. 1225–1233, 2014.
[25] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial intelligence in precision cardiovascular medicine," *J. Amer. College Cardiol.*, vol. 69, no. 21, pp. 2657–2664, 2017.
[26] R. L. Page, J. A. Joglar, M. A. Caldwell, H. Calkins, J. B. Conti, B. J. Deal, and J. H. Indik "2015 ACC/AHA/HRS guideline for the management of adult patients with supraventricular tachycardia: Executive summary: A report of the american college of cardiology/american heart association task force on clinical practice guidelines and the heart Rhythm society," *J. Amer. College Cardiol.*, vol. 67, no. 13, pp. 1575–1623, 2016.

CHUNYAN GUO received the master's degree from Inner Mongolia Medical University, in 2012. She is currently an Associate Professor with The Affiliated Hospital of Inner Mongolia Medical University. Her research interest includes cerebral protection in surgery.

JIABING ZHANG received the master's degree from PLA General Hospital, in 2019. He is currently pursuing the Ph.D. degree. He is also a Pharmacist-in-Charge. His research direction is biomedical engineering.

YANG LIU graduated from Inner Mongolia Agricultural University, in 2017. She is currently pursuing the Ph.D. degree. Since 2009, she has been a Docimaster with the Department of Clinical Laboratory, The Affiliated Hospital of Inner Mongolia Medical University. She is also an Associate Chief Technician. Her research interest is clinical basic examination research.

YAYING XIE graduated from Inner Mongolia Medical University, in 2002. She is currently pursuing the master's degree. She worked at The Affiliated Hospital of Inner Mongolia Medical University. She is also a Professor. Her research interest includes effects of anesthesia on organ function.

ZHIQIANG HAN graduated from Inner Mongolia Medical University, in 2002. He is currently pursuing the master's degree. He worked at The Affiliated Hospital of Inner Mongolia Medical University. He is also a Professor. His research interest includes effects of anesthesia on organ function.

JIANSHE YU graduated from China Medical University, in 1991. He is currently pursuing the master's degree. He worked at The Affiliated Hospital of Inner Mongolia Medical University. He is also a Professor. His research interests include effects of anesthesia and ABO blood group on organ function.