A Hybrid Approach for Identification of Coronary Heart Disease using PSO-DEFS

J.Kowsalya Devi* and Dr.P.Geetha**
*Research Scholar, Dept. of CSE, College of Engineering Guindy, Anna University, Chennai-600 025, India.
*jkowsalyadevi@gmail.com
**Associate Professor, Dept. of CSE, College of Engineering Guindy, Anna University, Chennai-600 025, India.
**geetha_akp@cs.annauniv.edu

Abstract: Coronary Heart disease (CHD) is a major cause of morbidity and mortality in the modern society. Medical diagnosis is an complicated task that should be performed accurately and efficiently. This study analyzes the Behavioral Risk Factor Surveillance System, survey to test whether self-reported cardiovascular disease rates are higher. Cardiovascular disease (CVD) can be classified into (1) chest pain (2) stroke and (3) heart attack. Heart care study specifies 13 attributes to predict the morbidity. Besides regular attributes other attributes such as BMI (Body Mass Index), physician supply, age, ethnicity, education, income, and others are used for prediction. An automated system for medical diagnosis would enhance medical care and reduce costs. Many existing methods found out the CHD affected patients by estimating the plaque thickness. These methods have the issues of lower accuracy and higher time complexity. In order to overcome the above limitations, hybrid approach based CHD detection is introduced in the proposed work. At first, the input image is preprocessed using green channel extraction and median filter. Subsequently, the features are extracted by gradient based features like HOG with CLBP. The texture features are concentrated with various rotations to calculate the edges. We present a hybrid feature selection which combines the PSO and DEFS for minimizing the time complexity. A binary SVM classifier categorizes the normal and abnormal images of patients. Finally, the patients affected by CHD are further classified by MLP. The effectiveness of the system is measured with the parameters such as accuracy, specificity and sensitivity.

Keywords: Coronary Heart Disease (CHD), Magnetic Resonance Image (MRI), Histogram of Oriented Gradient (HOG), Particle Swarm Optimization (PSO), Differential Evolution Feature Selection (DEFS).

Introduction

Coronary heart disease (CHD) is caused by hardening of artery walls due to cholesterol known as atherosclerosis. The disease progression is slow, asymptomatic, and may lead to sudden cardiac arrest, stroke, or myocardial infarction. Many imaging techniques are being employed to understand the classification of atherosclerotic plaques to estimate the risk, but they lack the required resolution and sensitivity for detection. In this paper, the experimental observations and habits of persons for predicting the risk factors of CHD are considered. The identification of risk factors helps in analyzing the patients for additional intensive tests.

ATHEROSCLEROSIS is a systemic/chronic disease characterized by the accumulation of inflammatory cells and lipids in the inner lining of the arteries. It is a leading cardiovascular disease causing deaths worldwide. Atherosclerosis has a long asymptomatic phase with a subclinical incubation period ranging from 30 to 50 years. The physicians would like to assess the risk of the patients having a severe clinical event such as stroke or heart attack and predict the same if possible. Some of the major known risk factors that eventually lead to the development of atherosclerosis are as follows: 1) family history of premature coronary heart disease (CHD) or stoke in a first degree relative under the age of 60; 2) tobacco abuse; 3) type II diabetes; 4) high blood pressure; 5) left ventricular hypertrophy; 6) high triglycerides; 7) high low-density lipoprotein (LDL) cholesterol; 8) low high-density lipoprotein (HDL) cholesterol; and 9) high total cholesterol. A large number of factors influence the onset of atherosclerosis making it a difficult task for the physicians to diagnose in its early stages.

Establishing a diagnostic procedure for early detection of atherosclerosis disease is very important as any delay would increase the risk of serious complications or even disability. Determining the conditions (risk factors) predisposing the development of atherosclerosis can lead to tests for identifying the disease in its early stages. Though the clinical risk factors of CHD are identified, there is a need for additional understanding on the disease progression for effective management. Researchers have been focusing on techniques for quantifying atherosclerosis plaque morphology, composition, mechanical forces, etc., hoping for better patient screening procedures. Imaging of atherosclerotic plaques helps in both diagnosis and monitoring of the progression for future management. There are two modes of imaging atherosclerosis: 1) invasive and 2) noninvasive.

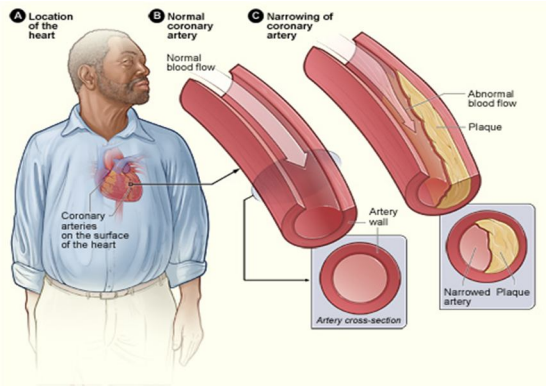

Fig.1 An illustration for diagnosis of Coronary Heart Disease (CHD), which leads to death

Among the invasive methods, X-ray angiography is the gold standard imaging technique even though it has certain limitations in providing information on plaque composition. The invasive procedures such as intravascular ultrasound and angioscopy help in understanding the plaque size and to a limited extent its composition. The intravascular thermography aids in monitoring the changes in plaque composition and metabolism. The noninvasive procedures such as B-mode ultrasound, computerized tomography (CT), and magnetic resonance imaging can provide information on plaque composition on vascular beds but they fail to throw much light on the metabolic activity of the plaque inflammatory cells. Though nuclear imaging techniques such as single-photon-emission-computed tomography and positron-emission tomography have the potential for 2-D and 3-D surface reconstruction of thrombus using radio labels to provide information on molecular, cellular, and metabolic activity of plaques, they lack the required resolution, sensitivity for detection, and functional assessment in medium- to small-size arteries found in coronary circulation.

Keeping in view the limitations of the imaging techniques, there is still a greater need for developing automated methods for predicting the risk factors of atherosclerosis disease in individuals which would be of great help in reducing the disease-related deaths. Machine-learning approaches have been employed in a variety of real-world problems to extract knowledge from data for predictive tasks. The presence of a large number of attributes in medical databases affects the decision making process as some of the factors may be redundant or irrelevant. Also, the presence of missing values (MVs) and highly skewed value distributions in the attributes of medical datasets require the development of new preprocessing strategies. Feature selection methods are aimed at identifying feature subsets to construct models that can best describe the dataset. The other advantages in using feature selection methods include identifying and removing of redundant/irrelevant features, reducing the dimensionality of the dataset, and improving the predictive capability of the classifier. This study attempts to identify risk factors causing atherosclerosis and the possible risk that the individuals are running at present.

The classification of plaque by American Heart Association (AHA)

- ✓ Type(I- II) Early Lesions – **Intimal thickening Intimal xanthoma**
- ✓ Type III Progressive lesion – **Pathological intimal thickening/ Pre atheroma**
- ✓ Type IV Progressive lesion – **Fibrous cap atheroma**
- ✓ Type V a, V b, V c Progressive lesion – **Healed cap rupture**
- ✓ Type VI Progressive lesion – **Thin fibrous cap atheroma**
- ✓ Type VII Progressive lesion – **Plaque hemorrhage and/or plaque rupture**

Many existing methods have the issues of lower accuracy and higher time complexity. Hybrid techniques are introduced to overcome these issues and the major contributions of the proposed work are:

- The green component is extracted using the green channel extraction and the noise present in the green component is eliminated by median filter.
- The contrast is enhanced using the histogram equalization.
- The features are extracted by Histogram of Gradient (HoG) along with Complete Local Binary Pattern (CLBP).
- A hybrid feature selection that incorporates the Particle Swarm Optimization (PSO) and Differential Evolution Feature Selection (DEFS) to reduce the time complexity.
- A binary Support Vector Machine (SVM) classifier categorizes the affected and not affected CHD. The abnormal images are then categorized the disease via the Multilayer Perceptron (MLP) classification.

To diagnose and follow up of Coronary Heart Disease (CHD), a better technique is employed for the clinical assessment. In the proposed system, the inclusion criteria are:

Table 1 Inclusion criteria for Coronary Heart Disease Detection

Sl. No	Attribute Name	Attribute Description	Attribute Values
1	AGE	Age in years	25-75 years
2	SEX	Male/Female	value 1: Male; value 0 : Female
3	CHEST PAIN	Chest Pain Type	value 1: typical type 1 angina, value 2: typical type angina, value 3: non-angina pain; value 4: asymptomatic
4	RESTBP	resting blood pressure	90-192
5	CHOLESTEROL	serum cholesterol in mg/dl	160-410
6	BLOOD SUGAR	fasting blood sugar > 120 mg/dl	value 1: > 120 mg/dl; value 0: < 120 mg/dl
7	ECG	resting electro cardio graphic results	value 0: normal; value 1: 1 having ST-T wave abnormality; value 2: showing probable or definite left ventricular hypertrophy
8	MAX HEART RATE	maximum heart rate achieved	71-202
9	ANGINA	exercise induced angina	value 1: yes; value 0: no
10	OLDPEAK	ST depression induced by exercise relative to rest	Continuous
11	STSLOPE	the slope of the peak exercise ST segment	value 1: unsloping; value 2: flat; value 3: downsloping
12	VESSELS	number of major vessels (0-3) colored by flourosopy	value 0 – 3
13	THAL	thalach state	value 3: normal; value 6: fixed defect; value 7: reversible defect

The rest of the paper is systematized as follows. Section I briefly overviews the literature review of detection and risk factors of Coronary Heart Disease (CHD). Section II involves the detailed explanation about the proposed methodology. Section III describes the results and discussion in detail. Section IV summarizes the performance analysis in terms of accuracy, sensitivity and specificity. Section V deals with a brief conclusive remark and discussion on future works.

Literature Review

This section reviews the relevant literature about the detection of coronary heart disease and its various techniques.

Author	Methodology	Description
K Srinivas et al	• Decision tree • Neural network • Bayesian network • Support Vector machine	Only 15 attributes were used. They achieved an accuracy of 89.7% and sensitivity of 90.17%
Mohamed Danesh Zand et al	• Multi layer perceptron network • ANFIS	Performance – 93%

Yan Zhang et al	• SVM	Three types of kernel functions namely linear, polynomial and RBF were used. They achieved an accuracy of 84.1%, 81.8%, 88.6% respectively. Lagging in feature extraction.
Sree Hari Rao et al	• Feature subset selection • Particle swarm Optimization • PRAA methodology	Accuracy – 98.04% Sensitivity – 93.75% Specificity – 100% • Lack in required resolution • Sensitivity for detection
Manju et al	• Multi layer feed forward neural network • Genetic algorithm	Accuracy – 85.3% Sensitivity – 83.5% • Various Optimization can be used
Jung Gi Yang et al	• ANFIS + LDA	Combined analysis technique to decrease uncertainty and increase accuracy. Only individual classifier
Hongyn Kang et al	• Logistic regression analysis • Statistical methods	Only one dataset is used
K Sudhakar et al	• Neural network • Decision tree • Naïve bayes	Associative classification
Chetana Yadav et al	• Association rule mining • Apriori algorithm	Accuracy – 93.75% Sensitivity – 95.65% • Deeper penetration in the diagnosis
Mahyar Taghizadeh Nouei et al	• ANFIS + K-means Clustering	Accuracy – 96.4% Different datasets are not considered
Sunitha Dodani et al	• Statistical Analysis (Mean and Standard deviation)	Both diabetic and non diabetic shows higher risk factor Lower sensitivity

Related Work

A Hybrid PSO-DEFS based Feature Selection for CORONARY HEART DISEASE Detection. This section presents the details of the hybrid approach for the detection of Coronary heart disease. Fig.2 depicts the overall flow diagram of the hybrid feature selection for identification of Coronary heart disease. The major components involved in the proposed method are:

- Pre-Processing
- Feature Extraction
- Feature Selection
- Classification

Preprocessing

To accomplish an image suitable for feature extraction, various steps are performed in preprocessing. The abnormalities associated with the input image are detected and have to be preprocessed in order to avoid the uneven illumination, to enhance contrast among the image background pixels and exudates, to eliminate noise in the input image. The techniques for preprocessing included in the proposed system are the green channel extraction, median filtering, and histogram equalization.

Green Channel Extraction

The heart image is taken in the form of RGB (Red, Green, and Blue) using MRI. The green channel of the RGB space is extracted and selected for identification of exudates. This is due to the most contrasted appearance of exudates in this channel. Therefore, the initial step is to transform this channel into a new image.

Median Filter and Histogram Equalization

The extracted green channel is then filtered using median filtering to eliminate the noise. Furthermore, histogram equalization is applied to the enhancement of contrast. The regions of exudates are much greater in intensity than the neighboring regions of the image. Therefore, the technique of image enhancement tends to assign them the greatest values of intensity.

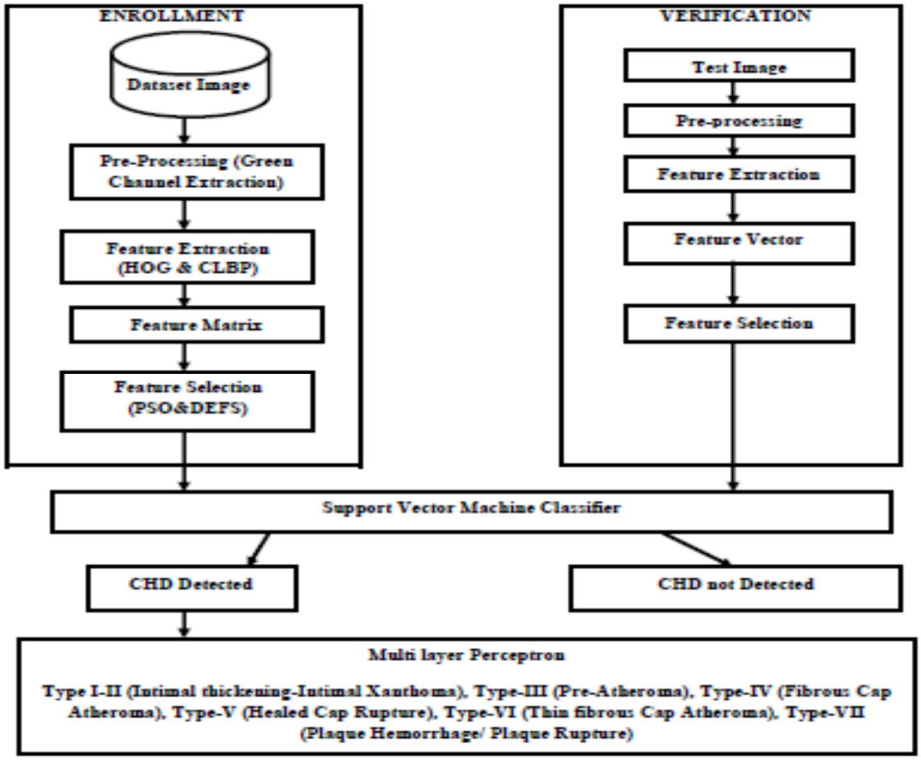

Fig.2 An Overall flow diagram of the proposed hybrid PSO-DEFS based feature selection for identification of Coronary Heart Disease

Feature Extraction

The preprocessed image is extracted using the Histogram of Gradient (HOG) and Complete Local Binary Pattern (CLBP). A set of features are defined to represent the information for analysis and classification. In the proposed feature extraction, 81 features are extracted based on the calculation of shape, position, contrast, brightness etc.

Histogram of Gradient (HOG)

HOG are the feature descriptors to detect the object and it counts occurrences of gradient orientation in the localized image portions. This process is same as the edge orientation histograms, contrasts of shape, and the descriptors of scale-invariant feature transform. As preprocessing provides a slight impact on the performance, the HOG ensures normalized color and gamma values by computing the gradient values.

Complete Local Binary Pattern (CLBP)

In CLBP, central pixel and Local Difference Sign-Magnitude Transform (LDSMT) represents a local region. After global thresholding, the central pixel is simply coded by a binary code and it is termed as the CLBP_Center (CLBP_C).The image local structure is decomposed into two complementary components by the LDSMT, difference signs and difference magnitude. Further, the two operators, CLBP-Sign (CLBP_S) and CLBP-Magnitude (CLBP_M) are employed to code them. All the three code maps, CLBP_C, CLBP_S and CLBP_M, are in binary format and so it can be readily combined to form the final CLBP histogram. The CLBP accomplishes much enhanced rotation invariant texture classification results than usual LBP based schemes.

Discussion of Genetic Algorithm (GA) vs. PSO and DEFS

The particle swarm optimization shares related characteristics to genetic algorithm, where the two algorithms traverse the search space by different fundamentals. The common elements shared by these algorithms are:

- In random manner, a population is initialized.

- An evaluation function is used to determine how fit a potential solution is.
- Based on fitness values, both algorithms are reproduction of the population.
- Both are generational for a predetermined amount of time.

The particles update the internal velocity as the PSO does not have genetic operators like crossover and mutation. In PSO, the mechanisms for information sharing are significantly different than the GA. Wherein, GA share information with each other by the chromosomes. Only the best neighborhood gives out the information to others in PSO, which is a one way information sharing method. The main benefits of the differential evolution consist in its simplicity than the GA technique. It has only three input parameters controlling the search process, namely, the size of population, the mutation and the crossover.

Feature Selection: PSO combined Differential Evolution Feature Selection (DEFS)

A population-based stochastic search algorithm is termed as the Particle Swarm Optimization (PSO), which uses swarm intelligence. Here, a particle denotes each member of the population and the swarm is the population. In the search space, every particle consists of a velocity and position. Each particle emphasizes its own best position, which is known as the individual best and it is denoted by pbest. Thus, the optimal position established by the whole swarm is called as the global best, meant by gbest. The entire particle has the value of fitness defined by the optimization function. At first, the position and velocities of each particle are randomly initialized. It finds the optimal position by updating their positions and velocities as under:

$$V_{p_I}(\tau + 1) = i_c V_{p_I}(\tau) + a_1 RN_1[pbest - P_I(\tau)] + a_2 RN_2[gbest - P_I(\tau)] \quad (1)$$

$$P_I(\tau + 1) = P_I(\tau) + V_{p_I}(\tau + 1) \quad (2)$$

where, V_{p_I} and P_I are the velocity and position of the particle, τ is the time, I is the index of the particle, i_c denotes inertial co-efficient that is employed for controlling the convergence behavior of PSO, and a_1 and a_2 are the learning factor.

For a selection of feature subset, a simple optimization Differential Evolution (DE) method is introduced. It is easy to utilize and has parallel and good convergence and for Fast implementation properties.

The following algorithm is the combination of PSO and DEFS algorithm.

PSO combined DEFS Algorithm

Inputs: Feature values, number of parameters to be optimized, maximum number of iterations

ALGORITHM: *Randomly initialize position, number of iterations, and the velocity of each particle*

```
While i< maxiteration
  Evaluate Fitness for each particle;
// Fitness
  Val=(feaval(i,j) – feaval(j,i))/mean(each attribute);
  If val>mean(attribute)/2
    pbest=val;
    For i=1 to popsize do
      Update pbest of particle;
      Update gbest of particle;
      For i=1 to PopulationSize do
                    For d=1 to Dimensionality do
          Update the velocity of particle i according to (2);
          Update the position of particle i according to (1);
      End For
    End For
  End For
End While
Return gbest;
gbest values passed to DEFS // attribute selected by PSO
For i=1 to size(inputs)
  Initialize population
  Generate chromosome // until number of iteration
  Process single point cross over
  Each chromosome undergoes mutation with a fixed probability μm.
  Evaluate by Fitness using:
      Fitness Value =∑ Euclidean distances of each feature in the column from their respective attribute.
End for
```

The selected features are then classified for the detection of Coronary Heart Disease.

Classification

The features obtained after feature selection are classified into normal and abnormal Coronary Heart Disease using Support Vector Machine (SVM) classifier.

Support Vector Machine (SVM)

SVM is the group of supervised learning tools applied to data classification and regression, which maps the training samples into various categories. These training samples are the points in feature space, whereas the test samples are also mapped to these similar feature spaces. It is then classified as belonging to any of the classes.SVM constructs a maximal splitting hyper plane among two classes. Therefore, the classification error is minimized. The input is mapped into a feature space of high dimensionality for linearly not separable data, where they can be divided by a hyper plane. Using kernel, this projection in high-dimensional feature space is effectively established. The SVM classifier with the optimum hyper plane is illustrated in Fig.3.

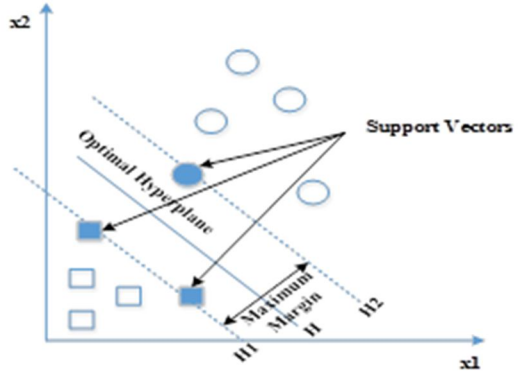

Fig.3 Illustration of SVM classifier with the optimum hyperplane

Multi-Layer Perceptron (MLP)

The resultant abnormal images are further classified by the multilayer perceptron classification. Based on the value of Plaque Thickness (PT) the diseases are classified. The affected CHD is identified as any one of the following type:

- Type(I- II) Early Lesions – **Intimal thickening Intimal Xanthoma**
- Type III Progressive lesion – **Pathological Intimal thickening/ Pre atheroma**
- Type IV Progressive lesion – **Fibrous cap atheroma**
- Type V a, V b, V c Progressive lesion – **Healed cap rupture**
- Type VI Progressive lesion – **Thin fibrous cap atheroma**
- Type VII Progressive lesion – **Plaque hemorrhage and/or plaque rupture**

Results and Discussion

This section analyses the results obtained from the investigation of the proposed methodology for the identification of coronary heart disease. The performance is evaluated based on the following metrics: Sensitivity, Specificity, Accuracy, and Receiver Operating Characteristics (ROC) classification. During the testing phase, 60 samples are taken and tested accordingly.

(a) (b) (c)

Fig.4 (a) The input heart image, (b) The preprocessed heart image, and (c) The histogram equalised heart image

All the collected datasets are real time datasets. The input heart image is collected, and then extracted from the green component and the noise is filtered.

Performance Analysis

Confusion Matrix – SVM Classification

Fig.5 shows the confusion matrix for the results obtained with SVM classification. The confusion matrix is determined among target class and output class. The diagonal values in green color denote the true positive and true negative values. Whereas, the diagonal values in red color indicate the false positive and false negative values. The confusion matrix exhibits the results of the true positive rate, false positive rate, accuracy, sensitivity, and specificity.

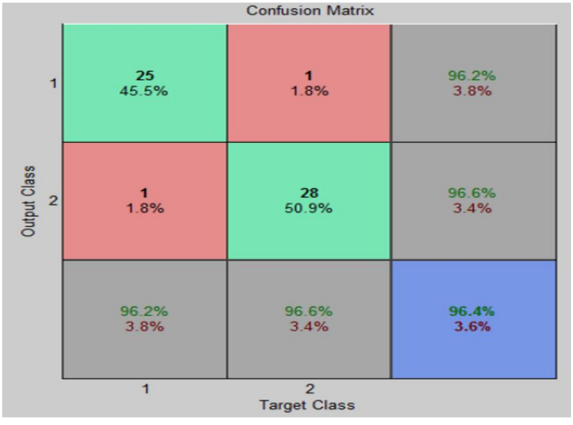

Fig.5 Confusion matrix – SVM classification

ROC Curve for SVM Classification

Receiver Operating Characteristic (ROC) graph for SVM classification is represented in Fig. 6. For predicting the accuracy rate, the graph is plotted between True Positive Rate (TPR) and False Positive Rate (FPR). ROC is formed by plotting TPR vs FPR at different thresholds. The proposed method results above 95% accuracy for SVM classification, hence the system performance is stable to detect the coronary heart disease. Also, the time taken for CHD detection consumes is less time. Hence, it is well suited for clinical applications.

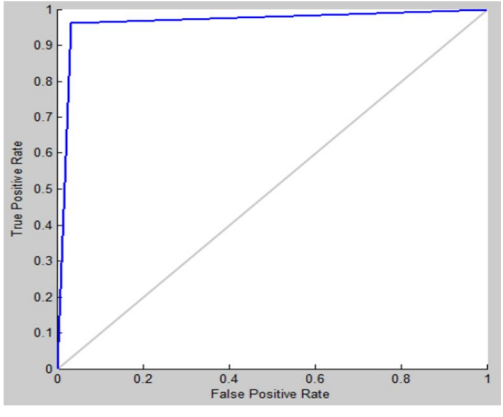

Fig.6 ROC curve graph for SVM classification

TABLE 2 depicts sensitivity and specificity calculation. It is found that the TP is 25, FP is 1, FN is 1 and TN is 28. The total test positive value is 26, total test negative value is 27, and the total population is 53.The sensitivity and specificity are obtained as 96.2 and 96.4 respectively as shown in the table.

Confusion Matrix for Multi-Layer Perceptron (MLP) Classification

Table 3 represents the confusion matrix for MLP classification. The diagonal values of the mentioned confusion matrix specify the number of inputs in each category.

Table 2 Calculation of Sensitivity and Specificity

	CHD disease present	CHD disease absent	
Test positive	True positives (TP)	False positives (FP)	Total test positives :
	25	1	26
Test negative	False negative (FN)	True negative (TN)	Total test negatives:
	1	28	27
	Total diseased: 26 Sensitivity =25/(25+1)	Total normal: 27 Specificity = 28/(1+28)	Total population: 53
	= 96.2	= 96.4	

Table 3 Confusion Matrix – Multi – Layer Perceptron

4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Accuracy

For the analysis of the accuracy, sensitivity, and specificity, the True Positive (TP) is the correct detection of number of images, False Positive (FP) is the incorrect detection of number of images, True Negative (TN) is the correct rejection of number of images, and False Negative (FN) is the incorrect rejection of number of images. Accuracy is defined as follows:

$$Accuracy = \frac{Number\ of\ TP + Number\ of\ TN}{Total} \quad (3)$$

Here,

$$Total\ population = Number\ of\ TP + Number\ of\ FP + Number\ of\ FN + Number\ of\ TN \quad (4)$$

Fig.7 shows the result of accuracy for the existing Neural Network (NN) classification and the proposed Multi-Layer Perceptron (MLP) classification.

Sensitivity

Sensitivity rate is described as the probability that a test result will be positive when the coronary heart disease is present. It is evaluated as follows,

$$Senisitivity = \frac{Number\ of\ TP}{Number\ of\ TP + Number\ of\ FN} \quad (5)$$

Fig.8 shows the result of sensitivity rate for the existing Neural Network (NN) classification and the proposed Multi-Layer Perceptron (MLP) classification.

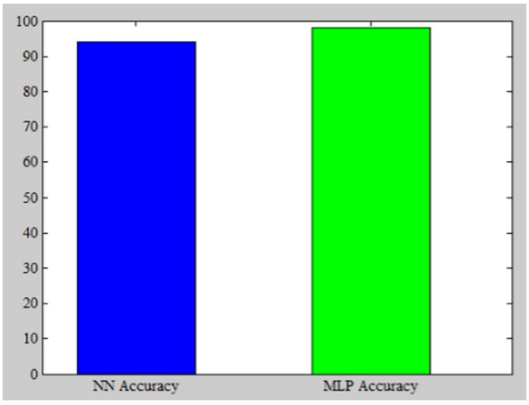

Fig.7 The result of accuracy for the existing NN and the proposed MLP classification

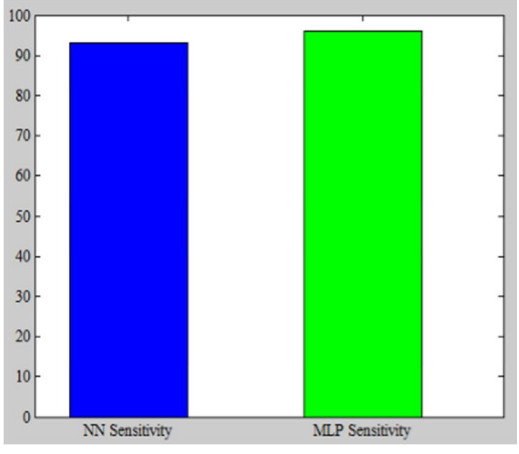

Fig.8 The result of accuracy for the existing NN and the proposed MLP classification

Specificity

Specificity rate is termed as the probability that a test result will be negative once the coronary heart disease is not present and is determined as,

$$Specificity = \frac{Number\ of\ TN}{Number\ of\ FP + Number\ of\ TN} \qquad (6)$$

Fig.9 shows the result of specificity for the existing Neural Network (NN) classification and the proposed Multi-Layer Perceptron (MLP) classification.

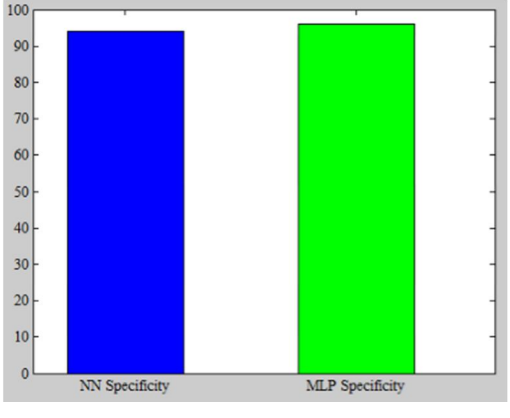

Fig.9 The result of specificity for the existing NN and the proposed MLP classification

Conclusion and Future Work

The foremost reason of fatality in the world is Coronary Heart Disease (CHD). This research work overcomes the limitation of time complexity and accuracy. The input image is preprocessed using the green channel extraction and the corresponding noise is removed by median filter. Using HoG and CLBP, the features are extracted and concentrated with different rotations for estimating the edges. In our scheme, a hybrid feature selection, namely, PSO and DEFS is employed to minimize the time complexity. A binary SVM classifier classifies the normal and abnormal images of CHD. The performance analysis of MLP classification shows better accuracy, sensitivity, and specificity compared to the existing neural network methods.

References

[1] [K.Srinivas, G.Ragavendra Rao, A.Govardhan " Analysis of Coronary Heart Disease and Prediction of Heart Attack in Coal mining regions using Data mining Techniques" IEEE fifth International Conference on Computer Science & Education , August 2010,pp 1344-349

[2] Mohammad Danesh Zand, Amir Hossein Ansari, Caro Lucas, Reza Aghaee Zade Zoroofi, " Risk Assessment of Coronary Arteries Heart Disease based on neuro fuzzy classifiers", Proceedings of 17th Iranian Conference of Biomedical Engineering, November 2010

[3] Yan Zhang, Figui Liu, Zhigang Zhao, Dandan Li, Xiaoyan Zhou, Jingyuan Wang, " Studies on application of Support Vector machine in diagnose of coronary heart disease" Sixth International Conference on ICEF, IEEE June 2012

[4] V. Sree Hari Rao, *M. Naresh Kumar* "Novel Approaches for Predicting Risk Factors of Atherosclerosis" IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS, VOL. 17, NO. 1, JANUARY 2013, pp 183-189

[5] T.Manju, K.Priya, R.Chitra, "Heart disease prediction system using weight optimized neural network", International Journal of Computer Science and management research, VOL.2 No.5, May 2013, pp-23912397

[6] Jung-Gi Yang, Jae-Kwon Kim, Un-Gu Kang, "Coronary heart disease optimization system on adaptive network based fuzzy inference system and linear discriminant analysis" Pers Ubiquitous Computing, Springer-Verlag September 2013

[7] Hongyu Kang, Lei An, Yi Xin, Xiaoming Hu, Qin Li, Ling Yin, Siyue Zhang " Prevalence and risk factors associated with chronic disease in new rural cooperative medical system: County in central China" Proceedings of CCIS2014 ,pp .319-322

[8] K.Sudhakar, M.Manimekalai, "Study of heart disease prediction using data mining", International journal of advanced research in computer science and software engineering, January 2014,pp. 1157-1160

[9] Preeti Gupta, Bikrampal Kaur, "Accuracy enhancement of heart disease diagnosis system using neural network and genetic algorithm" International journal of advanced research in computer science and software engineering, August 2014,pp. 160-166

[10] Chetana Yadav,Srikant Lade, Manish K Suman, "Predictive analysis for the diagnosis of coronary artery disease using association rule mining" International Journal of Computer Applications, VOL. 87, NO. 4, February2014, pp 9-13

[11] Mahyar Taghizadeh Nouei, Ali Vahidian Kamyad, Mahmood Reza Sarzaeem, Somayeh Ghazalbash, " Fuzzy risk assessment of mortality after coronary surgery using combination of adaptive neuro fuzzy inference system and K-means Clustering", Expert Systems, VOL.33 No.3, June 2016, pp-230-238

[12] Sunita Dodani, Gyanendra K Sharma, "Presence of coronary artery disease in diabetic and non diabetic south asian immigrants" Indian Heart Journal, Elsevier, 2017

[13] Vivek Reddy, Wael Taha, Shankar Kundumadam, Mazhar Khan, "Atrial Fibrillation and hyperthyroidism : A literature review" Indian Heart Journal, Elsevier, 2017

[14] H. Liu and R. Setiono, "A probabilistic approach to feature selection—A filter solution," in *Proc. 13th Int. Conf. Mach. Learn.*, 1996, pp. 319–327.

[15] A. Halder, *et al.*, "Dynamic Image Segmentation using Fuzzy C-Means based Genetic Algorithm," *International Journal of Computer Applications*, vol. 28,no. 6, pp. 15-20, 2011.

[16] Bharathi P T and S. P, "Differential Evolution and Genetic Algorithm Based Feature Subset Selection for Recognition of River Ice Types," *Journal of Theoretical & Applied Information Technology*, vol. 67, no. 1, 2014.