WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Machine Learning Based Heart Disease Prediction System Using CNN Algorithm

[1*] P.B. Dimsha, [2]Mrs. G. Krishnaveni

[1]Master of Computer Applications, Gnanamani College of Technology, Namakkal.
[2]Assistant Professor,Master of Computer Applications, Gnanamani College of Technology, Namakkal.
** Email Id: pbdimsha@gmail.com*

ABSTRACT

Nowadays heart disease is most dangerous to create heart disease suddenly attain heart attack leads to death. By increasing fast foods, genetic life cycle and age factors are the important reason for this disease which affects the coronary heart valves to block the blood flow. The development of medical science and engineering develops computer aided design based on image analysis with support of angiogram scans to identify the disease. Most of the prevailing techniques in machine learning models are failed to identify the disease properties in terms of feature analysis to identify the exact region of the block. Due to unidentified segmentation region and higher intensity, the identification accuracy is low due to higher false positives.

By addressing the problems, to improve the heart disease prediction using support vector quantization feature selection (SVQFS) based on Resnet50 convolution neural network (Resnet50-CNN). Initially the preprocessing is carried out by Min max -Neighbor vector normalization (Min-Max-NVN) to remove the noise and verify the actual margin in heart disease dataset. The heart disease impact margin rate (CSIR) is analyzed by identifying the actual affected features with support of mean covariance scalar estimation. Using the support vector quantization algorithm (SVQA), to find the actual features to select the important margins to reduce the feature scaling and dimension of non-scaling features. Then resnet50 –CNN is applied to predict the feature scaling with soft-max activation function and effectively predict the disease class to categorize normal and abnormal. The proposed system effectively segment the heart disease region by identifying the relevance feature margins and support vector dimension probability despond on disease impact rate to improve the detection accuracy. Also the performance increasing higher recall, precision rate by combative best rue positives feature limits with low false rate compared to the existing system.

Keywords: Heart Disease Prediction, Support Vector Quantization, deep learning, Convolutional Neural Network, Heart disease Impact rate, preprocessing and Normalization Machine Learning, Feature Extraction and classification

1. Introduction

In recent years, heart disease has emerged as a significant health concern, with its potential to cause sudden heart diseases and heart attacks leading to fatalities [1]. This alarming trend can be attributed to several factors including the rising consumption of fast food, genetic predispositions, and the inevitability of aging. The cumulative impact of these factors often results in the blockage of coronary arteries, which impairs blood flow and increases the risk of severe cardiovascular events [2]. To combat this pressing issue, advances in medical science and engineering have paved the way for computerized solutions that employ image analysis, particularly through the support of angiogram scans, to detect cardiovascular diseases [3]. However, prevailing methodologies, especially those that utilize machine learning models, frequently fall short of accurately identifying the properties of these diseases [4]. Issues such as unidentified heart disease properties in the feature margins and elevated intensity levels often lead to low identification accuracy and a high incidence of false positives [5]. Despite these advancements, numerous conventional machine learning models have encountered difficulties in accurately pinpointing disease characteristics, particularly in the nuanced feature analysis necessary for identifying specific regions of blockage. The current approaches often suffer from limitations in segmentation accuracy and are prone to generating higher rates of false positives, which undermine their effectiveness.

This paper addresses these challenges by proposing a robust methodology based on feature selection and classification using DL model. To enhance the accuracy of heart disease prediction through the implementation of Support Vector Quantization Feature Selection (SVQFS) in conjunction with a ResNet50 Convolutional Neural Network (ResNet50-CNN). The proposed approach begins with a comprehensive preprocessing phase utilizing Min-Max Neighbor Vector Normalization (Min-Max-NVN) to mitigate noise, thereby establishing clearer delineation within the heart disease dataset. Subsequently, the methodology employs a novel analysis framework called Heart disease Impact Margin Rate (CSIR), which aids in the identification of critical features through the utilization of mean covariance scalar estimation. This allows for the determination of relevant characteristics essential for selecting significant margins and reducing dimensionality—particularly in the context of non-scaling features.

Upon the identification of pertinent features, the ResNet50-CNN architecture is deployed to predict the likelihood of feature scaling utilizing a softmax activation function, thereby facilitating an effective classification of the disease states into normal and abnormal categories. The proposed system demonstrates an ability to accurately segment the heart disease-affected regions by identifying relevant feature margins and applying support vector dimension probability in alignment with the disease impact rate. Significant improvements in detection accuracy have been observed, with noted enhancements in recall and precision metrics, thereby achieving a reduction in false positive rates relative to existing methodologies. Ultimately, this research contributes valuable insights into the integration of advanced computational techniques for the effective diagnosis and predictive analysis of heart disease, with the potential to improve patient outcomes considerably.

2. Background study

The treatment and prevention of heart disease hinge on timely and accurate diagnosis. Traditional methods of heart disease detection [6], while effective to a degree, often do not leverage the full capabilities of contemporary machine learning methods like decision tree, KNN, naïve Bayes and ANN techniques [7]. Existing systems frequently tackle with significant challenges including the inadequacy of feature identification and poor feature analysis to find the region of affected areas [8]. This failure results in the inability to effectively localize the actual properties of blockage, leading to further complications in treatment and patient management. Addressing these shortcomings forms the occurred in identifying non related features leads low accuracy [10].

Cardiovascular diseases, including heart disease and heart disease, are leading causes of death worldwide, because of improper feature margins and dependencies get affected during analysis [11]. Traditional methods of Random Forest (RF), Convolution Neural Network (CNN) takes feature analysis for disease detection. However, with the advancements in machine learning and deep learning Support Vectors Machine (SVM) give more feature dimensionality problem solving approach [12, 13], researchers have explored the potential of these techniques to enhance the accuracy and efficiency of cardiovascular disease prediction [14, 15].

One study conducted a comprehensive review under artificial neural network (ANN) one of mostly covered model literature on the application of machine learning [16] and deep learning to heart failure diagnosis, readmission, and mortality prediction [17]. The researchers found that machine learning models have been extensively used to predict the presence of heart failure [18], estimate its severity, and forecast adverse events such as destabilizations, hospitalizations, and mortality [19]. These models have shown promising results, with the potential to improve disease management strategies, enhance patient outcomes, and reduce associated medical costs.

Similarly, another study investigated the decision tree approach in machine learning techniques is mostly used for absolute feature margin based classification and cardiovascular disease [20]. The researchers found that machine learning models can be valuable in identifying hidden patterns in various risk factors, which may not be easily discernible using traditional statistical methods.

2.1 Challenges in Existing Systems

Existing systems largely rely on classical feature extraction methods which do not adequately address the complexity and variability inherent in the data presented in clinical dataset. The inability to accurately identify the proper feature , disease property margins, coupled with the high intensity actual true positive rate, has led to diminished reliability in identifying disease properties. Moreover, the challenges of high-dimensional feature spaces often cause a decrease in model interpretability and increase the computational burden.

3.1 Proposed System Overview

To enhance the accuracy of heart disease prediction resulting from heart disease, we propose a novel approach that integrates Support Vector Quantization Feature Selection (SVQFS) with a ResNet50 Convolutional Neural Network (ResNet50-CNN). This system aims to improve the detection of disease properties by utilizing advanced feature analysis techniques.

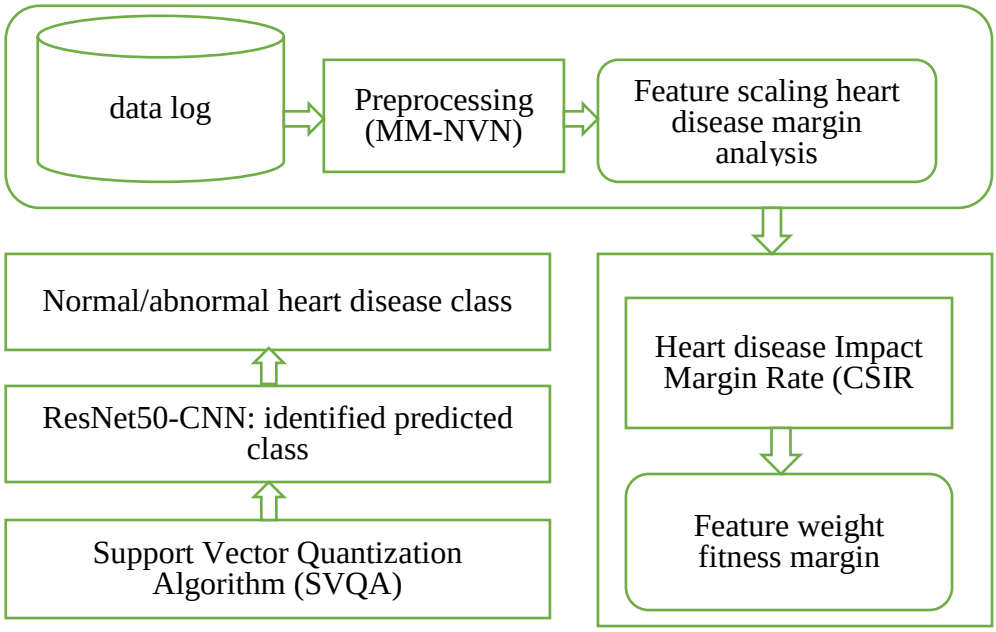

Figure 1: proposed architecture diagram-SVQA-RSCNN

The potential aspects of disease prediction is based on the feature margins, in similar the decision accuracy depends on the feature selection and neural classification. Figure 1explains the proposed architecture diagram-SVQA-RSCNN. By actively identify the feature margin to improve the classification accuracy. The deep learning CNN model enhanced with resnet50 dense layers to improve the prediction accuracy.

A) Preprocessing Phase

The initial step in our proposed system involves robust preprocessing of the disease dataset. We adopt Min-Max Neighbor Vector Normalization (Min-Max-NVN) to eliminate noise in the dataset and accurately delineate the margins concerning heart disease. This normalize the feature margin range, by formulating the actual value difference to reduce the improper feature dimension.

Step 1 : initialize the feature dataset set f (x,y)

Step 2:estimate scaking factor and margins $S_1(x,y) = \sum_{i=1}^{N} |x_i - y_i|$

Step 3: compute the difference $D(x,x') \geq \max_{(x'',y'')} D(x,x'')$

Hueristic diference $h(x) = \text{mode}(\{y'': (x'',y'') \in S_x\})$

Step 4:Calculate the Euclidean distance between the training data of the KNN model and the test data of hrt heart disease feature limits. The Euclidean distance matrix indicates the closeness of data points to each other within the K feature limits.

$$d_i = \sqrt{\sum_{i=1}^{p} (x_{2i} - x_{1i})^2}$$

The ka attains high accuracy by comparing the distance between training data and test data to determine the nearest point to reduce the noise to normalize the dataset.

B) Analysis of Heart disease Impact Margin Rate (CSIR)

Following preprocessing, we evaluate the Heart disease Impact Margin Rate (CSIR), which involves the identification of genuinely affected features through mean covariance scalar estimation. By analyzing these features, we can more accurately gauge the impact of various risk factors associated with heart disease the feature area analyzed by actual discriminative model which means that it computes $S(b|a)$ by identifying different possible values of class b based on a given input a. The equation to calculate this is below.

$$S(d|a) = d_i \sum_{x=1}^{R} u_x . f_x$$

The value of $S(b|a)$ cannot be directly calculated using the formula above. As a result, the values go from $-\infty$ to ∞ because you are not getting an output between 0 and 1 values. If the value is between 0 and 1, then the following exp function is used:

$S(d|a) = \frac{1}{v} exp \sum_x u_x . f_x$ to normalizing the feature limits $S(d|a) = \frac{exp(\sum_{x=1}^{R} u_x . f_x)}{\sum_d exp(\sum_{x=1}^{R} u_x . f_x)}$

The disease properties in impact rate a and output class candidate d. So, instead of f_x or $f_x(a)$, use $f_x(d,a)$ which assigns feature x of class d as specified input of a. By gettingprobability given b, a belongs to the d class:

$S(d|a) = \frac{exp(\sum_{x=1}^{R} u_x \cdot f_x(d,a))}{\sum_{d' \epsilon D} exp(\sum_{x=1}^{R} u_x \cdot f_x(d',a))}$

The maximum feature limits in disease relation get by argument, $\hat{u} = argmax_u \sum_b logS(b^y|a^y)$

The max marginal features limits to find the optimal model parameters that maximize the likelihood of heart disease data and provides a reliable classification to predict the presence of heart disease based on input features.

C) Support Vector Quantization Algorithm (SVQA)

Utilizing the Support Vector Quantization Algorithm (SVQA), we will identify and select key features that significantly contribute to the disease's manifestation. Through this process, we aim to reduce unnecessary data dimensions—specifically those related to non-scaling features thereby streamlining the dataset for more efficient processing. The support vector machine model is analyzed to optimize the minimum distance between two data points and determine the separating hyperplane. Furthermore, SVM techniques can be applied to analyze various data sets using different kernel functions, including radial deviation and linear functions. In binary classification, SVM can identify the maximum marginal hyperplane of training set samples through structural risk reduction analysis.

Calculate the maximum function margin of the average vector of the result hyperplane as shown in equation 1. Let's assume G-margin, $Arg_{\min}$ −minimum argument, f-decision hyperplane, z-normal vector, j-bias vector, u-values, a-data point.

$$G = S(d|a) \rightarrow \underset{u \epsilon f}{\text{Arg}_{\min}} d(u)$$

Calculate the number of samples in the feature space of the binary output feature vector of the training set dimension, as shown in Equation .

$$G = \underset{u \epsilon f}{\text{Arg}_{\min}} \frac{|u.z+j|}{\sqrt{\sum_{a=1}^{f} z_a^2}}$$

The classification function of SVM computes the sample's bias vector and the weight vector, as indicated in equation. Let's assume u_a −input feature vector, $u_a\{-1,+1\}$ −bnary output, m-number of samples in feature space, P^h −dimensional real space, z-weight vector, b-bias, K_q −training set.

$$K_q=\{(u_1,v_1),\dots,(u_g,v_h)\} \epsilon (P^h \times \{-1,+1\})^h$$
$$u_a \epsilon ([u_a]_1,[u_a]_2,\dots,[u_a]_h)^{Kq}$$

$$d(u) = z^{k_q}.u + j$$

$$d(u) = \begin{cases} z^{k_q}.u_a + j \geq -1 \; u_a \\ z^{k_q}.u_a + j \leq +1 \; u_a \end{cases}$$

The computation for solving the quadratic optimization problem using a rough edge SVM model is illustrated in equation 2. Maximizing the margin between two classes determines the distance between two hyperplanes, leading to the optimal hyperplane prediction. Let's assume G-minimum value, d-feature, T-Kernal function, ‖z‖ −hyperplane optimization.

$$\begin{cases} G(\text{in } z, j) & \frac{1}{2}\|z\| \\ \text{Subject to: } v_a(\langle z.u_a \rangle + j) & \text{for } a = 1,2,\dots,g \end{cases}$$

Above Equations demonstrate that the Lagrange multiplier for the binomial optimization problem is determined by analyzing the linear classifier in the binomial problem using the SVM model.

$$\underset{d,\xi_a}{G} \|d\|_T^2 + e\sum_{a=1}^{k} \xi_a$$
$$v_a d(u_a) \geq 1 - \xi_a \quad \text{for a } \xi_a \geq 0$$

Where Ma −maximim value, O-lagrange, e- Regularization parameter, a-slack variable, $u_a.v_b$ −represent measurement.

$$Ma\ O(\alpha) = \sum_{a=1}^{g} \alpha_a - \frac{1}{2}\sum_{b=1}^{g}\sum_{a=1}^{g} v_a v_b \alpha_a \alpha_b T(u_a.v_b)$$
$$\text{subject to: } \sum_{a=1}^{g} v_b \alpha_a = 0, 0 \leq \alpha_b \leq e, 1 \leq a \leq g$$

$$\underset{i_a}{G} \sum_{a=1}^{1} i_a - \frac{1}{2}\sum_{a=1}^{1}\sum_{b=1}^{1} i_a i_b v_a v_b \, t(u_a u_b) \;\; 0 \leq i_a \leq e, \text{ for a; } \sum_{a=1}^{1} i_a v_a = 0$$

The quantization value in SVM represents estimating smooth edges using Lagrangian variables while ensuring the shortest distance between the data and the hyperplane within predetermined boundaries. Let's assume q_T −shortest distance, T-distance.

$$G(O) = {}^{1}/_{2}\, z'z - \sum \lambda_T (\lambda_T (z'^{u_T} + j) + Q_T - 1) + i\sum Q_T \quad \text{for} 0 \leq i_a \leq e \text{ for all } i_a$$

Equation 11 depicts the kernel function and parameters for evaluating security and privacy functions through constructing the SVM model. Let's assume Q −security, d-security feature.

$$d_{(u,\alpha,j)} = \{\pm 1\} = Q\left(\sum_{a=1}^{g} y_a \alpha_a \, T(u, u_a) + j\right)$$

The SVM model chooses kernel functions and parameters to provide accurate results. For binary classification, especially in high-dimensional datasets, security and privacy aspects can be evaluated by constructing an SVM model. Therefore, optimal features can be achieved by selecting different feature subsets, estimating the impact rate of cardiovascular disease, and introducing a set of slack variables to assess heart disease data

D) Classification using ResNet50-CNN

The heart of our system lies in the implementation of the ResNet50 Convolutional Neural Network (ResNet50-CNN). This advanced neural network architecture is renowned for its deep learning capabilities, particularly in image recognition tasks. The CNN is employed to forecast feature scaling while employing the softmax activation function for effective classification into normal and abnormal categories. This stage of the system is crucial as it directly impacts the predictive values like diastolic rate, systolic rate, C-reactive proteins, thalac rate, VLDl, HDL, ECG, PCsd features take from dataset. The CNN method is one of the most successful applications and mainly works with transform layers, pooling layers and fully connected layers. Also, common layers using different layers of activation functions can significantly reduce the network training time. In addition, it can improve the generalization ability of the network and help prevent the overfitting problem of dropout layers. CNN models are considered hierarchical models, and the raw data can be extracted through transformation layers. The pooling layer can repeat the convolution and pooling operations with the dimension-reduced data as input to the lower convolutional layer. In addition, the high-level semantic information of the original data can be gradually transferred from the lower layers to the final layer. An integrated model can calculate the difference between the actual and predicted values. The CNN model presented in these can also be obtained by updating the relevant parameters together with a backpropagation algorithm.

By intent Resnet dense unit convolution kernel and impulse of input features. Let's assume c and d-features, $c_{j-A+1,y-B+1}$ −input features, E and F-highest convolutional kernel, T-parameter data, T_{AB} − convolutional kernel, j and i-multiple channels, A and B-convolutional layer.

$$d_{ji} = \mathrm{d}_{(\mathrm{u},\alpha,\mathrm{j})} \sum_{A=1}^{E} \sum_{B=1}^{F} T_{AB}\, c_{j-A+1,y-B+1}$$

Compute the convolution kernel involving multiple channels,Let's assume v-deviation, N- convolutional kernel size, $\mathbb{P}$ -filling parameter, q-layer.

$$w^{j+1} = [w^{q} \oplus z^{q+1}](j,i) + v = \sum_{T=1}^{T_j} \sum_{c=1}^{N} \sum_{d=1}^{N} [w_T^q Z_T^{B+1}(c,d)] + v$$

$$Q_{q+1} = \frac{Q_1 + 2\mathbb{P} - f}{h_{\circledcirc}} + 1$$

As described in Equation 37, estimate a function by assuming linear transformation in the fully connected function. Let's assume z-threshold weight.

Process Resnet50 layer unit→$d = z_c$

Calculate the error between the predicted value and the actual value as shown in Equation 38. Let's assume M-Error, M_e −estimated value.

$$M_e = soft\ ma \frac{\left|\sum_{j=1}^{F-1} (y_j - \circledcirc_j)^2\right|}{F-1}$$

Calculate the one-sample error rate as shown in Equation. Where M_x − error percentage.

$$M_x = \frac{\circledcirc_j - y_j}{y_j} \times 100\%$$

Estimate the normalization process as shown in equation

$$y_j = \alpha \frac{y_j - y_{min} + \beta}{y_{max} - y_{min} + \beta}$$

To Compute the modified CNN model determines the values by coefficients. Let's assume v-backpropagation, K-threshold, I-weight, M-error function, AJ − iteratively accuracy.

$$A_{(T+1)} = \bar{\omega}(T) - [I^k I + AJ]^{-1} I^k M$$

$$v_{(T+1)} = v(T) - [I^K I + AJ]^{-1} I^K M$$

Equation 43 is shown to evaluate the objective function to improve the output feature margin cognitive outcomes. Let's assume $N_{(w)}$ −obtained identification, M-error value, q-connection layer.

$$N_{(w)} = \frac{1}{1 + M^q}$$

In this category, feature relevance margin leverages objective functions to accomplish actual margin of disease class with CNN, Pooling Layers, and fully connected layers to proceeds to predict the output class.

4. Experiment evaluation

The efficacy of our proposed system will be measured through various performance metrics, such as recall and precision rates. We anticipate higher recall and precision compared to existing systems, owing to the significant reduction of false positives through our enhanced feature selection methods. This improvement will be crucial in informing clinicians and guiding patient treatment strategies, thereby reducing the risks associated with heart disease. The proposed system demonstrates improved performance compared to other existing systems. The results show higher accuracy, in precision, recall, F1-measure, and lower false rates in the detection of premature influences. Additionally, the implementation exhibits reduced time complexity, making it a viable solution for real-time applications.

Table 1: performance analysis in precision rate

Validation attained in precision rate %			
Methods/data's	500	1000	1500
Random forest	77.5	80.3	85.1
Decision tree	79.5	82.7	87.9
CNN	83.3	86.1	89.5
SVQA-RCNN	86.9	89.7	94.2

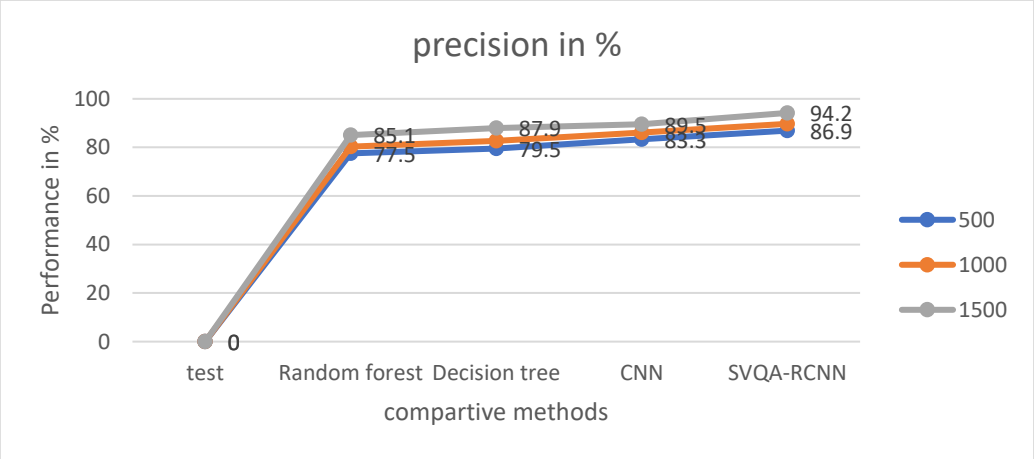

Figure 2: Analysis in precision rate

The SVQA-RCNN method's performance is tested in terms of precision accuracy by varying amounts of samples to validate the test results. Figure 2 explains the precision Analysis rate. The effectiveness of each class's tactics is evaluated and contrasted with the outcomes of proposed system is 94.2% compared to the other system. This increase the actual true positive rate and improve the prediction performance.

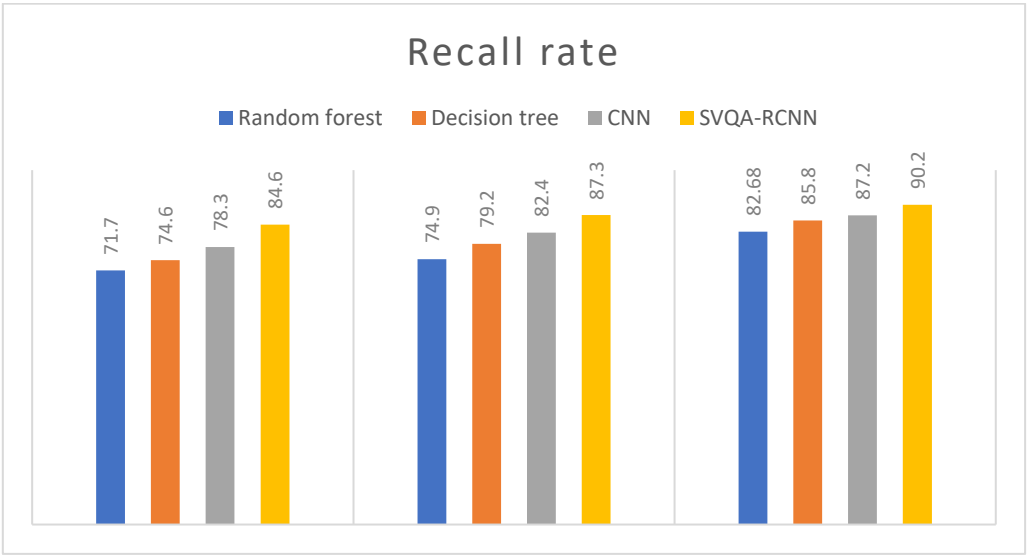

Figure 3: analysis of recall rate

In Figure 3, explains the performance of the recall rate by validating false positives rate which the proposed SVQA-RCNN system attains 90.2 % high and compared with different approaches to produce best performance than other methods. The result gained by accrual margins are relatively select by heart disease related features to improve the prediction accuracy.

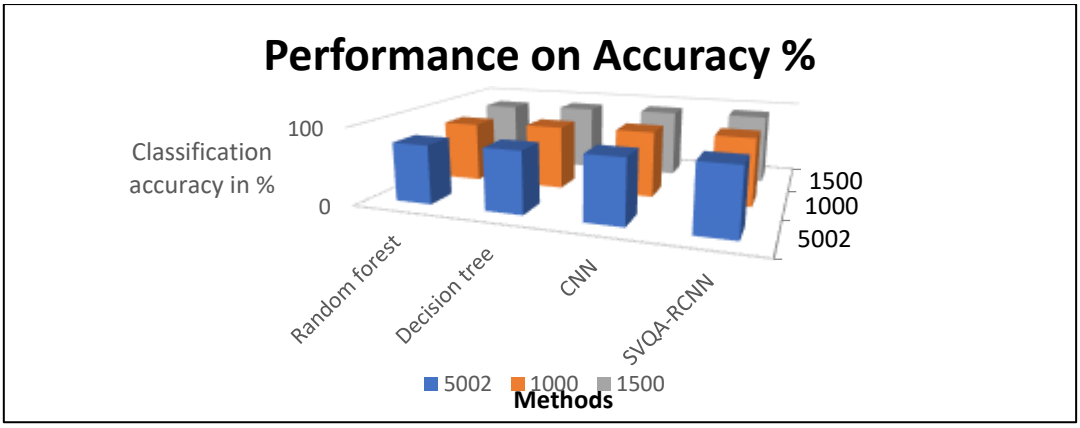

Figure 4: Analysis on classification accuracy

In Figure 4, we measure and compare the performance of classification accuracy. The classification accuracy of SVQA-RCNN method is 91.66% which is higher than other methods.

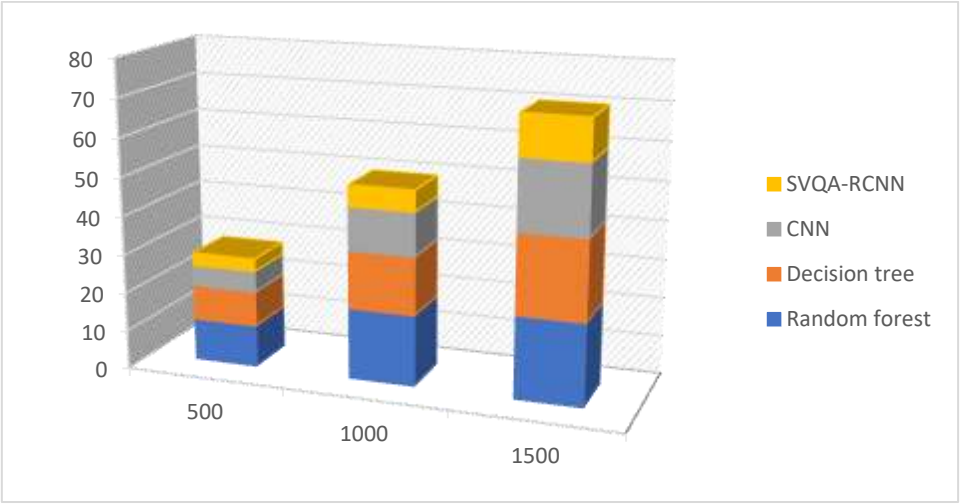

Figure 5: Analysis of Time Complexity

Numerous methods exist for measuring time complexity performance. Figure 5 shows the Analysis of Time Complexity Comparing the suggested SVQA-RCNN algorithm to other straightforward Bayesian, RF, and CNN techniques, time complexity is decreased. Compared to other approaches, SVQA-RCNN has a time complexity of 10.5 seconds, which is lower.

5. Conclusion

In summary, the proposed system leverages a robust framework that integrates SVQFS with ResNet50-CNN to address the prevalent challenges in heart disease prediction arising from heart disease. By focusing on accurate preprocessing by min max normalization, impactful feature selection support vector machine, and sophisticated classification of deep CNN techniques, our approach aims to improve the overall detection and segmentation accuracy. The anticipated SVQA-RCNN outcomes of this system improves the higher detection rate up to 95.2 % precision 94.6 %, as well in significant reduction in false positive rates, enhancing the reliability of cardiovascular disease assessments. Ultimately, as heart disease continues to pose a formidable health threat, innovations such as this proposed system represent critical steps toward harnessing technology in the pursuit of improved patient outcomes. Through continual refinement to adapt to emerging data trends, our system has the potential to make a meaningful impact in the realm of cardiovascular health.

Acknowledgement: Nil

References

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

2. .Songhee Cheon, Jungyoon Kim and Jihye Lim. (2019). "The use of deep learning to predict heart disease patient mortality", *International journal of environmental research and public health*, vol. 16, no. 11, pp. 1876.

3. K M, M. R, S. F. F and M. R. B. (2023). "Cardiovascular Heart disease Prediction System using Machine Learning Techniques," *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Coimbatore, India, pp. 89-93, doi: 10.1109/ICACCS57279.2023.10112727.

4. M. C. Das *et al*. (2023). "A comparative study of machine learning approaches for heart heart disease prediction," *2023 International Conference on Smart Applications, Communications and Networking (SmartNets)*, Istanbul, Turkiye, pp. 1-6, doi: 10.1109/SmartNets58706.2023.10216049.

5. E. Haque, M. Paul, S. Uddin, F. Tohidi and A. Khanom. (2023). "Analysis and Prediction of Heart Heart disease Using Lstm Deep Learning Approach," 2023 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Port Macquarie, Australia, pp. 340-347, doi: 10.1109/DICTA60407.2023.00054.

6. D. S Chaitrali and A.S Sulabha. (2012). "A Data Mining Approach for Prediction of Heart Disease Using Neural Networks", International Journal of Computer Engineering and Technology, vol. 3, pp. 30-40.

7. Wang Ren, Bian Di and Wang Shuqi. (2019). "Application of decision tree algorithm in heart disease risk classification prediction [J]", *Chinese Rehabilitation Medicine*, vol. 28, no. 3, pp. 233-236.

8. Zhu Qianli. (2019). "Application of artificial intelligence in the prediction of atrial fibrillation in heart disease patients [J]", *Hospital Management Forum*, vol. 36, no. 7, pp. 30-33.

9. M. N. Kumar, K. V. S. Koushik and K. Deepak. (Mar. 2018). "Prediction of Heart Diseases Using Data Mining and Machine Learning Algorithms and Tools", International Journal of Scientific Research in Computer Science Engineering and Information Technology, vol. 3, no. 3, pp. 887.

10. S SACHDEVA, T BHATIA and A K VERMA. (2018). "GI S-based evolutionary optimized Gradient Boosted Decision Trees for forestfire susceptibility mapping [J]", *Natural Hazards*, pp. 1-20.

11. S. Mohan, C. Thirumalai and G. Srivastava. (2019). "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access, vol. 7, pp. 81542-81554.

12. Patil, P. B. (2022). Reddy, Prediction of Cardiovascular Diseases by Integrating Electrocardiogram (ECG) and Phonocardiogram (PCG) Multi-Modal Features using Hidden Semi Morkov Model. International Journal on Recent and Innovation Trends in Computing and Communication, 10(10), 32–44.

13. T. Ozaki, K. Kitazume, K. Nishimura and T. Morioka. (2020). "Examination of risk assessment method based on emergency transportation record of heat heart disease", *Journal of Japan Society of Civil Engineers Ser. G (Environmental Research)*, vol. 76, no. 5, pp. 451-459.

14. Y. -Y. Xu, P. -S. Yen, Y. -H. Lin, C. -Y. Lai, Y. -C. Chen and C. -K. Liu. (2023). "Machine Learning-Based Outcome Prediction of Acute Ischemic Heart disease After Endovascular Treatment," *2023 IEEE/ACIS 8th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*, Hochimin City, Vietnam, pp. 374-377, doi: 10.1109/BCD57833.2023.10466305.

15. P. Govindarajan, K. S. Ravichandran, S. Sundararajan and S. Sreeja. (2017). "Impact of modifiable and non-modifiable risk factors on the prediction of heart disease disease," *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, Tirunelveli, India, pp. 985-989, doi: 10.1109/ICOEI.2017.8300855.

16. H. Zhang and F. Lin. (2023). "On Early Prediction and Diagnosis of Heart Disease Based on Random Forest Algorithm," *2023 4th International Conference on Computer, Big Data and Artificial Intelligence (ICCBD+AI)*, Guiyang, China, pp. 519-523, doi: 10.1109/ICCBD-AI62252.2023.00095.

17. V. K. Gnanavel, R. Sukanya, H. Mercy, P. Jayalakshmi and S. Veeramalai. (2022). "GUI Base Prediction of Heart Heart disease Stages by Finding the Accuracy using Machine Learning Algorithm," *2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)*, Chennai, India, pp. 1-5, doi: 10.1109/ICDSAAI55433.2022.10028815.

18. R. M. D. Saputra, Y. Chairul, D. Riana, A. S. Hewiz and F. Aziz. (2023). "Heart disease Prediction Based on Random Forest with SMOTE," *2023 International Conference on Information Technology Research and Innovation (ICITRI)*, Jakarta, Indonesia, pp. 17-21, doi: 10.1109/ICITRI59340.2023.10249261.

19. N. Riz Kifli, H. Hidayat, Rahmawati, F. Putri Sukoco, A. Rahma Yuniarti and S. Rizal. (2022). "Brain Heart disease Classification using One Dimensional Convolutional Neural Network," *2022 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob)*, Bandung, Indonesia, pp. 1-6, doi: 10.1109/APWiMob56856.2022.10014207.

20. T. N. Deepthi, S. Sharmila, M. Swarna, M. Gouthami and C. L. Akshaya. (2023). "Prediction of Brain Heart disease in Human Beings using Machine Learning," *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, Tuticorin, India, pp. 1314-1320, doi: 10.1109/ICEARS56392.2023.10085128.