DETECTION OF HEART DISEASE USING CONVOLUTIONAL NEURAL NETWORK BASED INSTANCE SEGMENTATION

N. Venkata Rao, Professor, Sasi Institute of Technology and Engineering, Tadepalligudem, A.P, India

ABSTRACT:
Human body prioritizes the heart as the second most important organ after the brain. Any disruption in the heart ultimately leads to disruption of the entire body. Being the members of modern era, enormous changes are happening to us on a daily basis that impact human lives in one way or the other. A major disease among top five fatal diseases includes the heart disease which has been consuming lives worldwide. Heart disease detection has long been considered as a critical issue. Machine Learning (ML) techniques find its use in medical sciences in solving real health-related issues by early detection and treatment of various diseases. Hence in this paper Convolutional Neural Network based Instance Segmentation using Machine Technique is presented for detection of Heart Disease. This paper intends to detect heart disease using datasets through a convolutional neural network and machine learning classification models like Random Forest and Gaussian Naïve Bayes (NB). The performance evaluated based on the accuracy, precision, recall, and F1-score for each of the models.

KEYWORDS: Health care services (HCS), Heart disease, Machine Learning (ML), Neural network, Random forest Gaussian Naïve Bayes.

I. INTRODUCTION

Heart diseases are number one source of death all-inclusive, medically called cardiovascular diseases (CVD) or strokes. Heart disease (HD) defined as a common term for the rise of plaque inside the coronary arteries that could lead to heart attack. Heart diseases are severe events which are caused by blockage inside the heart arteries. HD is a disease of long period and progresses slowly. Actually, heart disease is getting widespread proportion day by day. As a result, HD is the leading cause of death globally. According to the report of World Health Organization (WHO), heart disease causes death of millions of peoples annually than any other disease.

There are different kinds of CVDs, but most prevalent type of heart disease is contracting or obstacle of coronary arteries; it happens progressively over time. Heart pains, tininess of breath, deadness, fatigue, and pain in chin spine, upper abdomen throat or back are particular of most frequent signs of heart failure. However, several control variables help us lower risk of heart failure, like control of BP, lowered cholesterol, abstinence from smoking and routine exercise.

Diagnosing an infection with heart disease requires specialized cardiologists with complicated procedures and tests to figure the accurate and efficient treatment. Heart disease can be prevented by early diagnosis, followed by healthy eating, exercising, and avoiding alcohol consumption. In undeveloped countries, patients suffering from heart disease are diagnosed with

a severe delay at times, or they are transported over long distances unnecessarily, with the increase in the cost of travel and treatment, which is a burden on them.

To successfully diagnose the patients into has a heart disorder or not, Machine learning has been involved with other sciences to solve real-life problems automatically. Nowadays, ML algorithms are used for heart disease diagnosis and detection. Various studies have demonstrated an automatic detection to diagnose the severity of heart diseases using different machine learning techniques such as integrating multiple classification algorithms and Boosting algorithms to build strong automated detection systems.

Diagnosis by means of Machine learning is useful in the lack of medical personnel in addition that it is cheap and fast and can help doctors in the diagnosis process more efficiently to save patient's lives in early stages. Many of the studies show that machine learning techniques have gained significantly high accuracies in classification-based problems. In recent years, machine learning-based classification methods are one of the most operational appraisement methods for the research community and real-world applications.

II. LITERATURE SURVEY

In the growing field of data science and medical care, the need for automated detection diagnostic systems is increasing. Data scientists have developed several models, which have helped aid in the field of medical care. In the research, seven models used which are Support Vector Machine, K-Nearest neighbor, Logistic Regression, Decision Tree, Naïve Bayes, Random Forest, Artificial Neural Network models and some of ensemble techniques with the heart disease (HD) dataset.

Apeksha Shah, Swati Ahirrao, Sharnil Pandya, Ketan Kotecha and Suresh Rathod et al presented a Smart Cardiac Framework for an Early detection of cardiac arrest condition and risk [1]. Vardhan Shorewala et. al presents an approach of Early Detection of Coronary Heart Disease using Ensemble Techniques for detecting heart disease at early stages [2].

Hana H. Alalawi, Manal S. Alsuwat, et. al presents Detection of Cardiovascular Disease using Machine Learning Classification Models [3] to detect heart disease. Prasadgouda B Patila, Dr. P. M. Mallikarjun Shastry et. al presents Heart attack detection based on mask region based convolutional neural network instance segmentation and hybrid classification using Machine Learning techniques [4]. Princy et al. present a solution in predicting cardiovascular diseases by using the six machine learning classification algorithms which are: Decision tree, K nearest neighbor, Logistic regression, Naive Bayes, Random Forest, and Support vector machine [5]. The models have been applied on the cardiovascular disease dataset.

Rahma Atallah, Amjed Al-Mousa et al presents Heart disease detection using machine learning majority voting ensemble method [6]. Najmu Nissa, Sanjay Jamwal, Shahid Mohammad et. al presents Early detection of cardiovascular disease using machine learning techniques an experimental study [7] for detecting CVD in early stage. Md. Razu Ahmed, S M Hasan Mahmud, Md Altab Hossin et al presents a cloud based four-tier architecture for early detection of heart disease with machine learning algorithms [8]. Shadman Nashif, Md. Rakib Raihan, Md. Rasedul Islam et al presents Heart disease detection by using machine learning algorithms and a Real-Time Cardiovascular Health Monitoring System to monitor patient health care [9].

In this experiment, the prediction performance of different classification algorithms has been evaluated using the Stat Log Heart Disease dataset provided by the UCI Machine Learning Repository [10]. In order to develop a heart disease classifier, a data mining algorithm was built for data gathering and for predictive modeling. Thousand CHD patient records were mined, and the authors used a Support Vector Machine (SVM), Artificial Neural Network (ANN), and a Decision Tree (DT) for the binary classification job. Syedahamin Pouriyeh et.al the principal goal of this paper is to make a comparison of different machine learning techniques on a small dataset. The dataset in this research has been taken from cleave land database containing 303 instances with a total no. of 76 attributes among which a maximum of 14 attributes was used for the study. Due to lower variance in comparison with other estimators like single-fold approach, 10-fold cross-validation has been used for data portioning. Yang et al. proposed a multivariate regression model for the prediction of CVDs. The prediction was made on the collection of vital information of patients from the centers of the national high-risk program by having a cardiac events assessment. The dataset consists of all information, such as BP, cholesterol level, obesity, smoking, BMI, and many more attributes were taken into consideration.

III.HEART DISEASE DETECTION

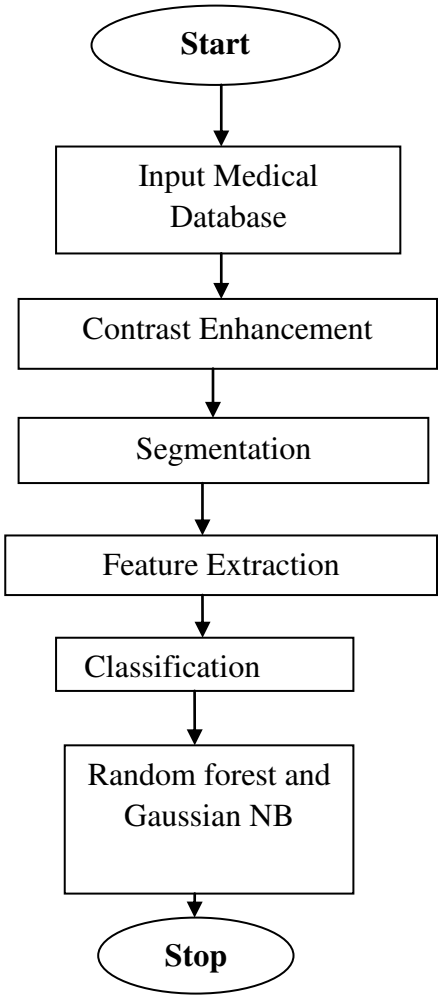

Fig. 1: FLOW CHART OF HEART DISEASE DETECTION

In this paper a Convolutional neural network based on image segmentation using machine learning techniques is presented for the detection of heart disease. The working flow of presented approach is shown in Figure 1. In this approach Ml algorithms like Random Forest and Gaussian Navie Bayes (NB) are utilized. The input data is taken from medical database which is given to preprocessing stage. The input can be arrhythmia 497, low ejection fraction 531, normal 617 is preprocessed by using resize the image and the contrast enhancement. In the segmentation stage the contrast-enhanced result is used for segmentation usinginstance segmentation then the output is given as an input for feature extraction stage. The segmented result is extracted to determine right feature set for classification, combinations of multiple heart-disease traits were subsequently checked. The image is trained and the classification using Random Forest and Gaussian Navies Bayes. Then finally detect the heart attack.

In image processing techniques, image resizing plays an important role in enlarging and decreasing specified image in pixel size format. Image interpolation is split into 2 distinct forms, which are image down sampling and up sampling, which are expected to fit either particular contact channel or output display when resizing data. Although sending low-resolution copies to client is more effective, an approximation of original high-resolution could be necessary for final visual data to be displayed. In several applications, precise resizing of image data is an important step, ranging from many consumer goods to vital functions in medical, safety and security sectors. Image segmentation is a process in which an image is divided into regions with certain homogeneous attributes, such as value of grey scale, color or texture.

Feature detection and extraction play a major role in computer vision field. Features are extracted from multi-scale images by using many algorithms. Features are extracted from original collection of calculated data and derived values are constructed. Improving quality and usefulness of research and classification is main objective of feature extraction. Classification is supervised learning method in which from input data, computer program studies and this learning is used to classify new observations in statistics and machine learning. Implementation of machine learning tool involves a wide variety of data science methods that are used to identify, classify and predict missing values.

3.1 Random Forest: It is a bootstrapping method with cart model. To construct cart model, consider 5 random samples and 100 observation samples and it built multiple trees with various initial variables. For each observation, same process is repeated 10 times to take final prediction which is each prediction function. This method is also called a supervised learning method. This method produces multiple random trees known as forests during learning phase. With highly correlated features, RF gives high predictive accuracy and is used in high dimensional issues basically in bioinformatics such as medical diagnosis.

3.2 Gaussian Naïve Bayes:

A Naive Bayes version that assumes normal distribution of Gaussian and accepts continuous data is Gaussian Naive Bayes. Naive Bayes Classifier is executed on high dimensional datasets. Naive Bayes extension is known as Gaussian Naïve Bayes. For estimating distribution of data, other functions may be used, but Gaussian is simplest to deal with so only need to approximate mean and SD from training data. An inference often taken when dealing with continuous data is that according to a normal (or Gaussian) distribution, continuous values associated with every class are distributed.

IV. RESULTS

For analyzing the performance of presented technique, the parameters like accuracy, precision, recall and F1- score are used. The measurement factors of these classified algorithms are calculated by the below terms: FP (False Positive), FN (False Negative), TP (True Positive) & TN (True Negative).

4.1 Accuracy: To predict presence and absence of heart attack, TP and TN measures classifier models' ability. Number of false predictions produced by models is identified by FP and FN

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

4.2 Precision: In heart attack classification model, it is a measure of sensitivity and success. It is defined as classifier probability when disease present gives result as positive which is called TP (True Positive) and calculated as

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

4.3 Recall: It is defined as classifier probability, when disease not presents gives result as negative which is called TN (True Negative) and calculated as

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

4.4 F1-Score: It is utilized to determine the prediction performance. It is defined as precision and recall weighted average. While 0 is worst, score value 1 is taken as best. It does not consider TNs. It can be calculated as

$$\text{F1-Score} = \frac{2*P*R}{P+R} \quad (4)$$

Comparison between the performance parameters of RF and Gaussian NB is shown in Table 1.

Table 1: parameters comparison between RF and Gaussian NB

Classifier	Accuracy	Precision	Recall	F1-Score
Random Forest	46.12%	48.28 %	46.44 %	43.62 %
Gaussian NB	55.34%	62.01 %	54.72 %	54.61 %

The graphical representation of accuracy and precision is represented in Figure 2.

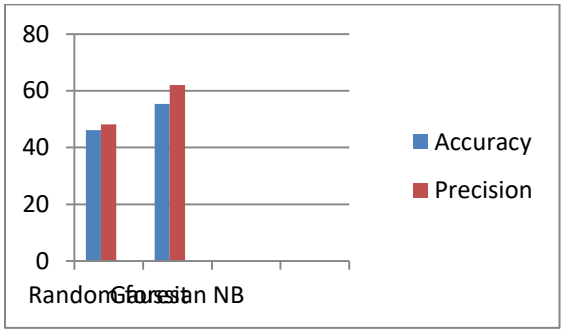

FIGURE 2: ACCURACY AND PRECISION COMPARISON BETWEEN RF and Gaussian NB

Figure 3 represents comparison between recall and F1-score.

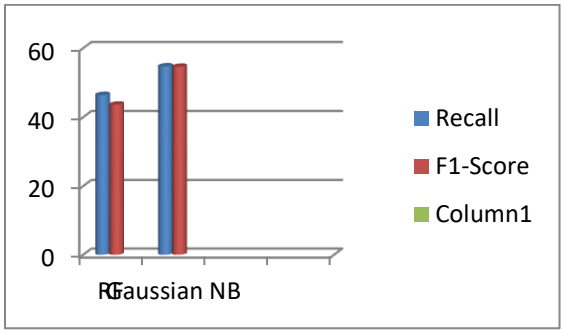

Figure 3: COMPARISON BETWEEN RECALL AND F1-SCORE PARAMETERS OF RF AND GAUSSIAN NB

V. CONCLUSION

The proportion of heart disease patients has been growing every day. There is a necessity for a framework that creates guidelines or identifies data by means of machine learning methods to solve this risky condition and worsen risks of heart failure disease. In this paper a Convolutional Neural network based instance segmentation using Machine Learning was presented for detection of heart disease. Presented method is based on machine learning classification which determines features extracted from clinical databases. Detection of heart attack using machine learning is effective and also experimental results show evaluated better results while using Random Forest and Gaussian Naïve bayes Machine learning classifiers.

VI. REFERENCES

1. Apeksha Shah, Swati Ahirrao, Sharnil Pandya, Ketan Kotecha and Suresh Rathod, "Smart Cardiac Framework for an Early Detection of Cardiac Arrest Condition and Risk", published: 22 October 2019 Frontiers in Public Health doi: 10.3389/fpubh.2021.762303

2. Vardhan Shorewala ,"Early Detection of Coronary Heart Disease using Ensemble Techniques", 2019 Published by Elsevier Ltd.
3. Hana H. Alalawi, Manal S. Alsuwat, "
Detection of Cardiovascular Disease using Machine Learning Classification Models",
International Journal of Engineering Research & Technology (IJERT),ISSN: 2278-0181, Vol. 10 Issue 07, July-2019
4. Hana H. Alalawi, Manal S. Alsuwat," Heart Attack Detection Based On Mask Region Based Convolutional Neural Network Instance Segmentation and Hybrid Classification Using Machine Learning Techniques", Turkish Journal of Computer and Mathematics Education Vol.12 No.9 (2019), 2228– 2244.
5. R. J. P. Princy, S. Parthasarathy, P. S. H. Jose, A. R. Lakshminarayanan, and S. Jeganathan, "Prediction of Cardiac Disease using Supervised Machine Learning Algorithms," in 2019 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2019, pp. 570–575.
6.Rahma Atallah, Amjed Al-Mousa, "Heart disease detection using machine learning majority voting ensemble method", 978-1-7281-2882-5/2019 IEEE
7. Najmu Nissa, Sanjay Jamwal, Shahid Mohammad, "Early detection of cardiovascular disease using machine learning techniques an experimental study", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-X, Issue-X, July 2019.
8. Md. Razu Ahmed, S M Hasan Mahmud, Md Altab Hossin, Hosney Jahan," A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms", 2018 IEEE 4th International Conference on Computer and Communications.
9. Shadman Nashif, Md. Rakib Raihan, Md. Rasedul Islam, Mohammad Hasan Imam," Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System", World Journal of Engineering and Technology, 2018, 6, 854-873.
10. X. Liu et al., "A Hybrid Classification System for Heart Disease Diagnosis Based on the RFRS Method.," Comput. Math. Methods Med., vol. 2017, pp. 8272091, 2017, doi: 10.1155/2017/8272091.
11. Weng, S.F., Reps, J., Kai, J., Garibaldi, J.M. and Qureshi, N. (2017) Can Machine-
Learning Improve Cardiovascular Risk Prediction Using Routine Clinical Data? PLoS ONE, 12, e0174944. https://doi.org/10.1371/journal.pone.0174944
12. Y. Yeh, C. Chen, C. W. Chiou and T. Chu, "A reliable feature selection algorithm for determining heartbeat case using weighted principal component analysis," 2016 International Conference on System Science and Engineering (ICSSE), Puli, 2016, pp. 1-4, doi: 10.1109/ICSSE.2016.7551594.
13. V. K. Dubey and A. K. Saxena, "Hybrid classification model of correlation-based feature selection and support vector machine," 2016 IEEE International Conference on Current Trends in Advanced Computing (ICCTAC), Bangalore, 2016, pp. 1-6doi:10.1109/ICCTAC.2016.756733

14. Singh, M., Martins, L.M., Joanis, P. and Mago, V.K. (2016) Building a Cardiovascular Disease Predictive Model Using Structural Equation Model and Fuzzy Cognitive Map. IEEE International Conference on Fuzzy Systems (FUZZ), Vancouver, 24-29 July 2016, 1377-1382.
15. Ghadge, P., Girme, V., Kokane, K. and Deshmukh, P. (2016) Intelligent Heart Attack Prediction System Using Big Data. International Journal of Recent Research in Mathematics Computer Science and Information Technology , 2, 73-77.