Journal of Theoretical and Applied Information Technology
15th May 2024. Vol.102. No 9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A HEART DISEASE DIAGNOSIS SYSTEM EMPLOYING RESIDUAL CONVOLUTIONAL NEURAL NETWORKS WITH ADAPTIVE CROSS-LAYER STACKED ARCHITECTURE

LAKSHMI ANUSHA KOTHAMASU[1], MAHANKALI SARITHA[1], B SWATHI[2], KOORAGAYALA SUKEERTHI[2]

[1, 2] Assistant Professor, Department of CSE, Vignan Institute of Technology and Science, Deshmukhi, Telangana-508284, India

E-mail: [1]anushakothamasu21@gmail.com

ABSTRACT

When all diseases are considered, heart disease is the leading cause of mortality worldwide. Heart disease is a significant risk. Whenever the blood and oxygen-supplying arteries to the heart become completely blocked or constricted, a cardiac issue results. Heart disease is one of the primary causes of death. In a short period of time, there has been a substantial increase in mortality. Accurate and secure diagnosis should be given priority in healthcare since a misdiagnosis of heart illness may result in mortality. If the prognosis is true, although cardiovascular disease may be avoided, if the prediction is incorrect, it may be dangerous. As a result, the team has created an ACLS-RCNN based ICSOA technique-based Intelligent heart disease prediction model. In the beginning, the patient data is acquired and pre-processed. Handling missing value handling, the three primary components of data pre-processing include scaling with IQR-RS, managing imbalance data using ROS, and pre-processing the data itself. Features are extracted using Kernel-based Linear Discriminate Analysis after the pre-processed data. Then, to eliminate pertinent components, Utilizing hybrid optimization, also known as (CI-AO+GI-SMO), valuable data is selected. Using recommended Cross-Layer Stacked Residual Convolutional Neural Networks that are Adaptive, an illness will be predicted. The prediction technique is improved by using the Improved Cuttlefish-Swarm Optimization Algorithm (ICSOA). Using Python simulation, the recall, accuracy, f1-score and precision of the suggested system are compared to those of other well-known methods. With regard to clinical data and diagnoses, our technology is capable of making exact predictions.

Keywords: *Improved Cuttlefish-Swarm Optimization, Convolutional Neural Network, Heart Disease, Correlation-based feature selection, Kernel-based Linear Discriminant Analysis*

1. INTRODUCTION

The use of deep learning (DL) is prevalent, including education and health. Due to improved PC accessibility and power, ubiquitous learning has been replaced with profound learning as innovation has developed. Medical services make use of profound comprehension in several ways [1]. To locate instances and develop predictions, the medical services industry can use images, patient data, and other data products. Further applications of profound learning can be discovered in the treatment of cardiac diseases. Any condition that affects the heart's normal blood-siphoning limit is referred to as coronary illness. The body's coagulation is impaired, which causes cardiovascular collapse. A healthy heart and a cardiac condition are contrasted in Figure 1.

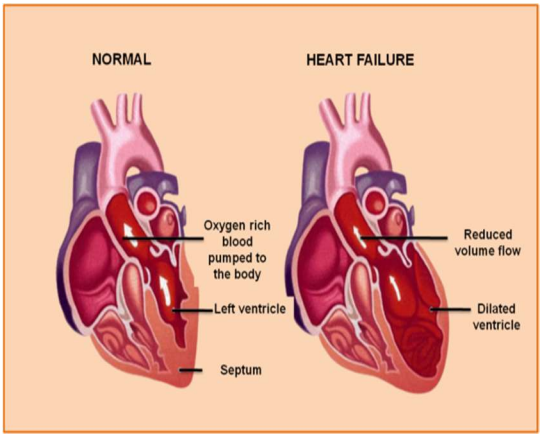


Figure 1. An example of a heart in good health and one that is failing

Another name for cardiovascular illness is cardiovascular sickness. Through clinical

investigation, a few risk factors for cardiovascular failures and coronary vein disease have been identified. There are two different categories of risk factors: things that are inconsistent with previous real realities are immutable, yet lifestyle changes like giving up smoking or enhancing health may help decrease cholesterol and blood pressure. By changing one's lifestyle and using prescribed medication, some of these risk factors could be lowered or eliminated [2]. Among the numerous clinical side effects of coronary disease, referred considered to be ischemic coronary disease on sometime, are myocardial localized necrosis (MI), angina pectoris, and unexpected cardiac arrest. The primary cause of these issues is coronary atherosclerosis, which are the outcome of complex interactions between the myocardium and coronary spread. [3]. Experts think that angiography is the most reliable approach for examining computer-aided design and utilize it the most frequently. However, it is additionally associated with a hefty cost and significant negative effects [4]. These problems call for painless methods to differentiate cardiovascular disease at this time. Traditional techniques for identifying cardiovascular infections depend on knowledge of a patient's history, any accompanying symptoms, and a formal evaluation report. Although they are not the particular ones, these are extremely important.

To determine the timing and possibility of a diagnosis of coronary disease, several studies have been carried out and various AI models have been applied. The congestive cardiovascular breakdown classifier developed by Melillo et al. distinguishes between individuals who are at high risk and those are protect. 7]; They utilized "Truck," an artificial intelligence computation, stands for "Relapse" and "Order". According to the results, they were 93.3 percent receptive and 63.5 percent explicit. The exhibition electrocardiogram (ECG) strategy, which utilises deep brain networks to choose the finest components and then apply them, is then put out by Rahhal et al. for future development [8]. The clinically appropriate emotionally supporting network is then provided by Guidi et al. in order to recognise cardiovascular breakdowns [9] and prevent them in their early phases. In addition to other AI and deep learning models, they sought to explore brain organisations, help vector machines, irregular backwoods, and truck computations. Truck and irregular woods outperformed everyone in the category with an accuracy of 87.6%. The standard-based technique and natural language processing are combined by Zhang et al. [10] The unstructured clinical notes were used to calculate the NYHA HF class with 93.37 percent accuracy. Using standard data including the patient's age, blood pressure, and glucose level, Parthiban and Srivatsa's SVM algorithms had an accuracy rate of 94.60 percent. Parthiban and Srivatsa [11] used them to recognise diabetes individuals and consequently foretell cardiovascular disease.

Data that is very dimensional is a prevalent issue in AI, as is the curse of dimensionality, which is the challenge presented by some of the datasets we use when attempting to display substantial amounts of data in three dimensions [12]. As a result, examining this data consumes a lot of RAM, and in rare circumstances, the data may develop dramatically on its own, leading to overfitting. The clear repetitiveness of the dataset may be reduced using the weighting components, which will enhance handling and execution time [13, 17]. To minimise the dimensionality of a dataset, a range of component designing and element selection procedures can be used[18]. Due to human error, these procedures frequently result in incorrect conclusions [5, 6]. It is essential to develop a computerised framework for diagnosing heart illness that takes taken into consideration various cross-layer stacked lasting convolutional brain organisation (RCNN for the upper leg tendons).

1.1. Commitments of this examination

By providing a clever clinical decision framework for recognising cardiac sickness in light of current mechanical deep learning methodologies, the examination research makes a contribution. In order to extract highlights, they utilise a part-based LDA (K-LDA) and half-breed streamlining (CI-AO+GI-SMO) to choose educational data. The crucial features are removed using relationship-based include choice (CFS). Utilising the readily available Kaggle cardiovascular illness assortments, our developed framework was made and evaluated. All of the handling and computations were done in Python.

The following sections make up the remainder of the article: Section 2 looks at recently published works. In Section 3, the suggested procedure is described. The results of the investigation are presented in Section 4. The conclusion of the investigation is given in Section 5.

2. RELATED WORKS

S. Mohan, G. Srivastava, C. Thirumalai, et al., [20] enhanced the precision of cardiovascular illness forecasts by developing a way for using AI algorithms for identification. The expectation model is illustrated using a number of widely used order mechanisms and element combinations. Considering the use of a direct model and a cross-breed irregular timberland, they were able to increase their exhibition level, with an accuracy level of 88.7%, for the heart illness model.

Fitriyani et. al. [21] created a useful expectation model for coronary disease. To adapt the information transmission and detect and eliminate any exceptions, this model utilises Thickness Based Spatial Bunching of utilises with Commotion, a modified nearest neighbour (Destroyed ENN), and XGBoost to forecast cardiovascular disease. The outcomes of the model were contrasted with those of other models, including logistic regression (LR), multi-facet perceptron (MLP), naive bayes (NB), support vector machine (SVM), random forest (RF) and decision tree (DT), as well as results from earlier reviews (Statlog and Cleveland), using two publicly available datasets. The resulting model achieved exactnesses, outperforming both rival models and prior review results. The Coronary disease CDSS (HDCDSS) model was also developed by them to help physicians and experts analyze the severity of the subjects' and patients cardiovascular disease in light of their current status. Therefore, early therapy may be started in order to prevent the fatalities brought on by late coronary disease study.

Khan and others [22] developed an Internet of Things (IoT) system using a Changed Profound Convolutional Brain Organisation (MDCNN) for a more accurate assessment of heart illness. The patient- monitored smart cardiac monitor keeps a close watch on the patient's electrocardiogram (ECG) and circulatory strain. The MDCNN is used to classify the received sensor data as either normal or unusual. The framework's presentation was examined for brain groupings and strategic relapse. The findings support paradigm methods based on heart disease expectation. The proposed technique shows that, for the largest quantity of data, the MDCNN outperforms existing classifiers with a precision of 98.2.

El Hassani and others [23] utilised a dimensionality reduction methodology and the component selection method to find signs of heart illness. The UCI AI Archive provided the data for this experiment, termed Coronary Illness. Six ML classifiers accepted the dataset's 74 highlights and name. The Cleveland dataset showed that chi-square plus head part analysis (CHI-PCA) with randomised woods (RF) had a 99.0 percent efficiency, which was only 98.5 percent higher than the next greatest accuracy. The physical and physiological parameters from the study that ChiSqSelector gathered included the heart vessels, chest discomfort, and highlights connected with ST gloom the experiment's results.

Atiqur Rahman and others [24] recognised the under- and overfitting of the prediction model as its drawbacks. By keeping the model from over- or underfitting, it may show outstanding performance on both datasets preparation data and testing data. Overfitting the preparation materials often results from improper organisation and unneeded components. To eliminate immaterial components, we advise utilising a two-dimensional model. To identify the optimal deep brain organisation (DNN), we use a rigorous search methodology. The proposed half-and-half model, 2 - DNN, was assessed by contrasting its models, one extra state-of-the-art AI model, and recently unveiled techniques for cardiac illness prediction. encouraging as compared to the prior approaches discussed. The results of the analysis imply that utilising the recommended suggestive framework, clinicians would be able to accurately predict heart illness.

Dutta and others [25] generated a layer-based all of the clinical data with class imbalances may be explained by a convolutional neural network. Cardiovascular disease (CHD) incidence is forecasted using the National Health and Nutrition Examination Survey (NHANES) data. With a class-explicit implementation that is reasonably consistent, our basic two-layer CNN adapts to the discomfort. Unlike the majority of AI models currently in use that use this type of information, which even when class-explicit loads are changed, withstand class irregularity. Reaching trendy 1 could get more challenging with additional test data (real CHD expected rate) accuracy and precision in a very unbalanced dataset. Using a completely related layer, the critical parts were first homogenized before moving on to the succeeding convolutional phases.

3. PROPOSED APPROACH

Among the most popular techniques for diagnosing cardiac issues, an angiogram is

considered to be one of the most reliable.. Angiograms have a wide range of drawbacks, such as a high price tag, numerous incidental effects, and a high amount of specialised information, to identify a few of them. Regular techniques frequently result in incorrect conclusions and take longer due to human error. It was difficult to identify the border and get rid of the shocking disturbance. It is very much demanding to get recurrence widths with low edge values for several cardiovascular conditions. The proposed method has been investigated using the flowchart below. Using this as a basis for development, they were able to understand the complex structure of the connection and how it functions.. In this approach, we'll use a number of cycles to spot coronary disease. By examining the patient's cardiac condition, this method helps avoid treating the patient later than necessary. As seen in Fig. 2, we should talk about the suggested method in the following steps.

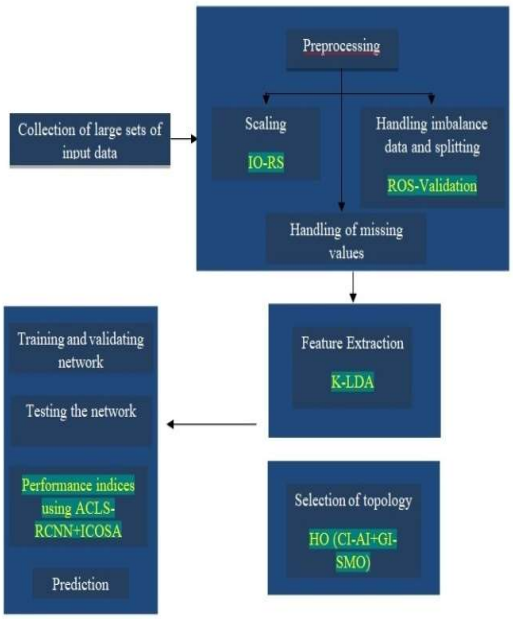


FIGURE 2. FLOW OF THE PROPOSED METHOD

3.1. Data Collection

The Cleveland, Cardiovascular Sickness Forecast, and UCI datasets from Kaggle are used to evaluate this strategy. We contrasted the idea with the theories of other authors. To test the sensitivity of the methods, we examined two different publicly available datasets. The cardiovascular arrhythmia dataset from the college of california irvine (UCI) Machine consists of electrocardiographic (ECG) signals from a small number of peoples that includes more than 300 distinct organic informational markers. Nonetheless, the distribution of cases is not really equal. 70,000 patient records from one Kaggle dataset were used to discuss cardiovascular disease. Serum full is included in this Kaggle dataset. Every patient has their cholesterol, Blood pressure readings, systolic and diastolic, myocardial status documented and proportional glucose level.

$$A_{Coll} = \left[A_1, A_2, A_3, A_4, A_{mis\,sing}, A_{23} \ldots\ldots A_n \right] \quad (1)$$

3.2. Pre-processing

The statistics gathered are inaccurately scaled, lacking, and unbalanced, among other things. The pre-processing phases include normalisation, redundant data reduction, and missing value imputation. The system is prepared for diagnostic after pre-processing. The doctor's dataset is then processed.

Unstructured data processing might result in erroneous node deployment and scheduling methods, which waste a lot of energy. To answer this challenge, the work tackles such problems and pre-processes data.

$$B^{\pm} = B^{pre}\left[A_{Coll}\right] \quad (2)$$

3.2.1. Handling of missing data

Consider the dataset as a data matrix, which is subsequently divided into define a missing value A_{Coll} to $A_1, A_2, A_3, A_4, A_{mis\,sing}, A_{23} \ldots\ldots A_n$, that show the observed and missing data. Let $A_{mis\,sing}$ represent a missing value framework that is characterised by,

$$A_{missing} = \{0, if\ A\ is\ \ observed 1, if\ A is\ missing\} \quad (3)$$

Let's Q define a vector of expected numbers that represents the correlation between the $A_{mis\,sing}$ missingness in and the total number A_{Coll} of observations. The likelihood of whether a value is different or absent relates to the way the missing quality systems are represented, as we shall see later.

Missing completely at random (MCAR)

This develops as the absence of knowledge is independent of both observable and unobservable

measurements. A clear definition for MCAR probability is provided:

$$P\left(\mathrm{A}_{missing} \mid Q\right) \quad (4)$$

Missing At Random (MAR)

The probabilities of missing values in MAR is exclusively linked to evidence that has been observed. The following factors determine the chance of MAR:

$$P\left(\mathrm{A}_{missing} \mid \mathrm{A}_{Coll} Q\right) \quad (5)$$

Data sets from research in the health sciences frequently contain missing at random (MAR). Observable predictors can fill in missing values in this method.

Missing Not At Random (MNAR)

When neither MCAR nor MAR are the missing piece of information, then this is true. The noted and absent values do not completely determine. This method's handling of missing characteristics is frequently challenging since it depends on new datasets. These steps are used to calculate the MNAR probability:

$$P\left(\mathrm{A}_{missing} \mid \mathrm{A}_{Coll}, \mathrm{A}_{missing}, Q\right) \quad (6)$$

A position's chance of being seen or missing based on the both

$\mathrm{A}_1, \mathrm{A}_2, \mathrm{A}_3, \mathrm{A}_4, \mathrm{A}_{missing}, \mathrm{A}_{23} \ldots\ldots \mathrm{A}_n$.

3.2.2. Handling imbalanced dataset

Active network identification is seriously impaired by the large amount of data that technological advances continually produce. The attributes of big data, which encompass data volume and transmission speed in addition to content, characterize this aspect. A dataset is considered huge data in terms of information volume when it is challenging to examine it using traditional analytical tools.

There are two classes in the imbalanced dataset: a class that is the majority and a class that is the minority, the former having a far larger sample sizes. The machine learning represents a majority class as a result of this differentiation, which disrupts the process of learning new information. The random frame interpolation method described by has been used to complete this task:

$$\mathrm{A}_{Coll} = \mathrm{X}^{ros}\left[sampling\_strategy = '\min ority'\right] \quad (7)$$

3.2.3. Scaling

Data can be standardised by scaling the independent characteristics that are present within a specified range. When dealing with information that has previously undergone radically shifting order of magnitude, values, or categories, it is utilised. A neural network model would often consider larger values to be higher and judge smaller values to be lower, regardless of the measuring instrument, if the procedure is not carried out. Data outliers might sometimes end in inaccurate data scaling.

Since it is necessary to exclude outliers from the calculation of the mean and normal deviation, input data must be standardized, the study uses a strong data scalability technique. A robust scaler (IQR-RS) based on the inter quartile range (IQR) is then used to scale the parameter using the calculated values.

Calculated to attain are the percentage distributions of the 50th percentile, 25th percentile, and 75th percentile. After that, the values of each variable are multiplied by either the standard deviation or the confidence interval (IQR) and the median are then removed, this is determined by either the confidence interval (CI) measurement or the difference in rank between the 75% and 25% percentiles.

$$\mathrm{A}_{Coll} = \frac{\left(\mathrm{A}_1 - median\left(\mathrm{A}_{coll}\right)\right)}{\left(p_{75} - p_{25}\right)} \quad (8)$$

The conversion of unstructured information to structured information is facilitated by taking into account the previously mentioned restrictions. As a consequence, topological selection is performed on the relational database.

3.3. Using Kernel-based Linear Discriminant Analysis (K-LDA) for feature extraction.

When inner period irregularities are confused and their presentation is approximated using hurriedly created test information, piece-based straight discriminant analysis is a straightforward solution. This method assures maximal detachability by raising the ratio between changes to within class variety in each informative index. In order to solve the grouping problem in photo acknowledgment, straight discriminant analysis is performed. To obtain more accurate results, we must devise methods that use LDA as opposed to head parts examination.

1. These are constructed using the optimum information assortments and preliminary lattices. To make things simpler to understand, let's represent the information sources as an information lattice using the accompanying arrangement.

$$\text{Set1}=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot \\ b_{n1} & b_{n2} \end{bmatrix} \quad (9)$$

$$\text{Set2}=\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ \cdot\cdot & \cdot\cdot \\ \cdot\cdot & \cdot\cdot \\ c_{n1} & c_{n2} \end{bmatrix} \quad (10)$$

2. For each data set, calculate the overall insensitivity and the cruel. The average of all the data gathered after the merger is q2, and the method for sets 1 and 2 is q1. For both the full information that was gathered by integrating, condition gives the standard.

$$\pi_3 = q_1 \times \pi_1 + q_2 \times \pi_2 \quad (11)$$

3. To determine the requirements for class distinctness, K-LDA uses both inside scattering. The predicted covariance of each individual within the class is scattering. The dissipate estimation frameworks are established using conditions 3 and 4.

$$R\_u = \sum\_i \ [\![q\_i \times (\ [\![cov]\!]\ \_i\ )\ ]\!] \quad (12)$$

$$R\_u = 0.5 \times \ [\![cov]\!]\ \_1 + 0.5 \times \ [\![cov]\!]\_2 \quad (13)$$

$$[\![cov]\!]\ \_i = (y\_i - \pi\_i\ )\ (y\_i - \pi\_i\ )\text{\^{}}D \quad (14)$$

$$R\_c = \sum\_i(\pi\_i - \pi\_3\ ) \times (\pi\_i - \pi\_3\ )\text{\^{}}D \quad (15)$$

Using criterion (7), the covariance network is done.

Every gathering's intermediate vectors are those that emphasize the differences between groups and their component elements. As previously mentioned, the ideal rules are in no way dissolved inside each scatter. The focal point of the altered space is the arrangement made feasible by the extension of these standards. Nonetheless, are employed to ascertain the most persuasive rationale for the class-subordinate shift. If, as is improbable, the LDA is only a classification, it requires a tailored upgrading of the standards. Considerations for streamlining that a class subordinate might have include the following:

$$criterion = inv(R_u) \times R_c \quad (16)$$

$$criterion_i = inv(cov_i) \times R_c \quad (17)$$

The cutoff thresholds imposed by Condition 10 on the middle lists of the gatherings suggest. Limits for grouping information, as seen in the photos, were stated by entirely switching the information's scope.

Transformed set = $transform\_spec^D \times data\_et^D$ (18)

$$transformed_{se\ _i} = transform\_i^D \times se \quad (19)$$

The test vectors are also modified and ordered using the estimation. These models illustrate grouping in a scenario with two classes. We present the original informative collections and the corresponding informative indexes generated after change. These differences demonstrate how change pushes the boundaries of classification accuracy. Transformation is used because it can be difficult to discover a good position where classes overlap in the picture space, even if the classes in this model were well-defined. Where to make the best adjustment is at the hub with the biggest Eigen vector. They are especially fascinating since they demonstrate how the highest degree of information can be viewed as the quadratic development.

3.4. Multiple objectives feature analysis

Hybrid optimization techniques are used to handle the problem of determining the optimal visual representation of the distribution of courses in entrepreneurship and innovation. The proposed hybrid technique combines the standard Aquila Optimizer with Spider Monkey Optimization, which is based on Glorot initialization. The hybrid optimization that has been developed takes care of the best stability between local and global rates. By striking a balance between the speeds of discovery and exploitation, the most significant and useful knowledge can be discovered for teaching courses on innovation and entrepreneurship. The algorithm focuses on eliminating the poorest results in order to obtain the best possible solution.

Originally developed as a population-based optimization method, the Aquila Optimizer was modelled after how an Aquila naturally interacts with its prey. Since random initialization of the upper and lower bounds results in a significant slowdown in the convergence rate of local solutions due to variance in local search solutions, confidence interval-based population initialization is used to overcome this problem. There are four categories in which the proposed Hybrid enhancement

calculation's advancement approach falls: CI-based population introduction and high takeoff and vertical stoop search space selection, utilising within a join search space by GI-SMO, investigating inside a veer search space by shape trip with explained it to float attack, and finally updating GI-SMO's new search area.

The starting population of newcomer accommodations is established in line with the confidence interval, which is determined by:

$$\varsigma = \begin{bmatrix} \varsigma_{1,1} & \cdots & \varsigma_{1,Dim} \\ \vdots & \cdots & \vdots \\ \varsigma_{N,1} & \varsigma_{N,J} & \varsigma_{N,Dim} \end{bmatrix} \quad (20)$$

$$ub_j = \bar{\varsigma}_{ij} + \Im \frac{\sigma}{\sqrt{n}} \quad (21)$$

$$\varsigma_{ij} = rand \times \left(ub_j - lb_j\right) + lb_j, i = 1,2,....Dim \quad (22)$$

$$lb_j = \bar{\varsigma}_{ij} - \Im \frac{\sigma}{\sqrt{n}} \quad (23)$$

Where $\bar{\varsigma}_{ij}$ shows the average candidate sample value, $\Im$ indicates the value of the confidence interval representing the popularity of a certain candidate among the same group of people, and $\frac{\sigma}{\sqrt{n}}$ displays the population's distribution of the neighbouring candidates, or the standard deviation.

The CI-AO finds the best hunting spot with a high takeoff and vertical stoop, all without skipping a beat. The assessment begins by considering appropriate actions to take when stooping upright during takeoff. According to research, this straight equals Eq.

$$\varsigma_i(t+1) = \varsigma_{best}(t) \times \left(1 - \frac{t}{T}\right) + \left(\varsigma_M(t) - \varsigma_{best}(t) * rand\right) \quad (24)$$

Where, The primary search technique generates the next cycle, which will be set up (ς_i), $\varsigma_{best}(t)$ is the established arrangement up until the focus that reflects the erroneous location of the prey.

After then, a quick float assault to the prey inside the Aquila rings completes the form trip. Here, the AO only scans the targeted area of the intended prey in order to get ready for an attack. The mathematical description of this behavior is found in Eqn.

$$\varsigma_2(t+1) = \varsigma_{Best}(t) \times Levy(D) + \varsigma_R(t) + (Y - X) * rand \quad (25)$$

Where, $\varsigma_2(t+1)$ in this search strategy is organised and routinely produced (ς_2). Using the aspect break in D, which displays the toll appropriation running away work Eq. (12). $\varsigma_R(t)$ is an unconventional decision taken in the setting of $[1, N]$ at the i th iteration

$$levy(D) = a \times \frac{u \times \ell}{|v|^{\frac{1}{2}}} \quad (26)$$

$$\begin{aligned} &where, \\ &r = r_1 + u \times d_1 \\ &\theta = -\omega \times d_1 + \theta_1 \\ &\theta_1 = \frac{3 \times \pi}{2} \end{aligned} \quad (27)$$

$$\begin{aligned} &\ell = \left(\frac{\kappa(1+\varepsilon) \times \sin\left(\frac{\pi\beta}{2}\right)}{\kappa\left(\frac{1+\varepsilon}{2}\right) \times \varepsilon \times 2^{\left(\frac{\beta-1}{2}\right)}} \right) \\ &Y = r \times \cos(\theta) \\ &X = r \times \sin(\theta) \end{aligned} \quad (28)$$

accepts a r_1 value somewhere within the range of a predefined value of 0.00565. d_1 is represented roughly set at 0.005 and indicates the whole number e-dimensional search space (Dim).

When the Aquila receives exact prey location information, it dives higher in preparation for an initial attack. To find out the prey's reaction, this is done. The Aquila's slow drop attack, low trip behaviour can be mathematically explained by Eq.

$$\varsigma_3(t+1) = \left(\varsigma_{Best}(t) - \varsigma_M(t)\right) \times \lambda - rand + \left(\left(ub - lb\right) \times rand + lb\right) \times \tau \quad (29)$$

Where, $\varsigma_3(t+1)$ provides the foundation for the subsequent cycles, a third the achievement strategy that offers (ς_3), $\varsigma_{Best}(t)$ refers to the prey's difficult environment till focus (the most advantageous configuration), and $\varsigma_M(t)$ indicates the average value of the current configuration at t^{th} using intensity as a measurement Eq. (15). $rand$ is a random value between 0 and 1, somewhere in that range, λ and τ are the boundaries the fixed-use double-dealing GI-SMO, lb determines the upper bound of the specified problem and connects to the inferior bound.

The spider monkeys' ingenious foraging techniques added intrigue to the GI-SMO process. The dividing combination social structure is the foundation for the spider monkeys' foraging strategy. The characteristics of a group are determined by its social association when a female leader chooses to divide or unite it. A Glorot instatement is now familiar with ways to avoid the previously mentioned problems as a result of the effort. Utilising the GISMO approach conducted, the exploitation changes in the CI-AO is finished.

After the research phase, the second part of the computation is double-dealing, weather it updates the arrangements in light of a choice chance that may be decided to be a component of wellbeing.

from the goal work f_t

$$f_t = \begin{cases} \dfrac{1}{1+f_i} & if\ f_i \geq 0 \\ 1+abs(f_i), & if\ f_i \leq 0 \end{cases} \quad (30)$$

The choice likelihood Pb_i is resolved on the roulette determining the wheel. Should it happen, if f_t is SM's health, the chance that it will be selected for the global pioneer stage is calculated using both of the following two formulas:

$$Pb_i = \frac{f_i}{\sum_{i=1}^{N} fitness_i} \quad (31)$$

In order to update the location, the information about the global pioneer, the know-how of the nearby SM, and SM's own determination are all used by SM. The following is the position update requirement for this stage:

$$\varsigma_{nemij} = \varsigma_{ij} + (ub_j(0,sd)) \times (\varsigma_{GLkj} - \varsigma_{ij}) + ub_j(-1,1) \times (\varsigma_{rj} - \varsigma_{ij})$$
$$sd = \sqrt{\frac{2}{\varsigma_{ij} + \varsigma_{GLkj}}} \quad (32)$$

Where sd demonstrates the startup's standard deviation provided by Glorot. Andis the location of the world's inventor in the j^{th} dimension aspect. The three components of this position update condition are as follows: the parent (current) SM's diligence is shown in the first segment, and her interest in the global pioneer is shown in the second; and the random nature of the calculation is preserved by the third component.

During the global pioneer learning step, the calculation observes the optimal answer for the entire crowd. It was believed that the legendary SM was the most significant figure in history. The location of the global pioneer is also checked, and if it hasn't been updated, the worldwide pioneer-connected counter, known as, is raised by one instead of being left at 0. The global leader limit (GLL) is compared with the Global Boundary Calculation for the Global Pioneer.

Phase of the Neighbourhood Leader Learning: In this part of the calculation, the group members' covetous determination is used to update the local pioneer's location. If the neighbourhood pioneer doesn't update its state, the local pioneer is assigned a number called local limit count (LLC), which is incremented by one before being reset to 1. The nearest leader for each gathering is identified using this method. The Local Leader Limit, which is a legitimate edge (LLL), is reached by increasing the Nearby Limit Count counter until it reaches.

The neighbourhood leader decision phase: Prior to this level, both the local as well as global pioneers are recognised. People from that group should update their viewpoints. Suppose if any of the local leaders are failed to do so, through arbitrary introduction or by reflecting on the experience of the global pioneer. Equation (33) is used to calculate the irritation rate.

$$\varsigma_{nemij} = \varsigma_{ij} + (ub_j(0,1)) \times (\varsigma_{Lj} - \varsigma_{ij}) + ub_j(0,1) \times (\varsigma_{rj} - \varsigma_{ij}) \quad (33)$$

Because the existing local pioneer has run out of resources and hasn't been updated to the LLL number of emphases, this group's intentions are directed toward the global pioneer to alter the current hunt bearings and positions. The situation serves as evidence for this. In the end, the inquiry bearings and locations let experimenting with several course regulator configurations until the optimal one is found.

The accuracy of learning is increased while dimensionality and noise are reduced through feature selection. The effectiveness and efficiency of many feature selection techniques are constrained by data dimensionality."Correlation-based Feature Selection" is a quick method. Using wrapper selection methods, FCBF finds k-related features from the quality dataset. The dataset is divided into P segments for each elimination phase, and the lowest-scoring feature is eliminated. One characteristic is removed by FCBFiP, a unique FCBF modification that looks for data duplication. Compared to FCBF, FCFBiP is more effective.

3.5.Prediction

Adaptive Cross-layer Stacked Residual is the highest quality CNN. In Figure 3, The architecture of the ACLS-RCNN is shown. The proposed method is much shallower than traditional CNNs, which are employed in medical image processing. From CT data, multilayer context data must be recovered, in the network, they add a shared and cropped (SPC) layer.

With improved accuracy and faster convergence, the new damage detection system gathers visual data at a deeper neural network level.

Pre-processed, fused data is utilized for testing and training.

$$F = \left[\Theta_{Coll} + I_{coll}\right] \tag{34}$$

The convolutional layer receives each sub-band and conducts between F_n^G, where G and n and indicate the feature indices and sub-band levels. Convolution layer calculations are done as

$$\hbar_n^G = \sum_{i=0}^{n}\sum_{j=0}^{n} \omega_{i,j}^G F_{i,j}^G + \zeta_{i,j}^G \tag{35}$$

$$\hbar_n^G = Z\left(\sum_{i=0}^{N}\sum_{j=0}^{N} \omega_{i,j}^G F_{i,j}^G + \zeta_{i,j}^G\right) \tag{36}$$

Where, i, j represents the characteristics as rows and columns., $\omega_{i,j}^G$ indicates the subband weight, suggests bias, and implies the $\zeta_{i,j}^G$ function of the RLU activation.

The convolutional image is used to both give and prevent object distortion while maintaining the size of the image at a consistent length. It helps to keep vision loss at bay. The final convolutional layer, which has three pyramid levels (1x1, 3x3, and 6x6) totalling 46bins and a fixed-length output, is then supplemented with the STP layer.

$$\hbar_n^G = \mathrm{P}_{layers}^{STP}\left(\left|F_{i,j}^G\right|\right) \tag{37}$$

Where, $\mathrm{P}_{layers}^{STP}$ a layer for STP pooling. Last but not least, the convolutional neural network layer is formed by the combination of the convolutional layer and the STP pooling layer., which may extract low-level information and simplify the image. Thus, the output of this convolutional layer is compressed and delayed.

$$\hbar_m = act_{sm}\left(\sum_{i=1}^{n} \omega_{i,j} F_{i,j} + \zeta_{i,j}\right) \tag{38}$$

$$\hbar_n^G(F_{Flatten}) = \left[F_1, F_2, F_3, F_4, \ldots . F_n\right] \tag{39}$$

The probability-based image detection is offered by the softmax. The ROC curve is used for assessing the threshold probability. The new approach requires a high true positive rate or lower. The 0.88 TPR and 0.12 FPR cut-off values were selected. The points outside the curve indicate the normal image, and the points inside the curve represent the damage image.

For CNN learning, chain rule and vector calculus are utilized.

Assume that is a vector β and that is a (i.e., $\beta \in R$) scalar. Thus, the partial β derivative with respect to be Ω vector, if is a function of Ω, is denoted by the following expression:

$$\left(\frac{\partial \beta}{\partial \Omega}\right)_i = \left(\frac{\partial \beta}{\partial \Omega}\right)_i \tag{40}$$

Its i-th element is $\begin{pmatrix}\partial\beta\\ \partial\Omega\end{pmatrix}_i$: Notably,

$$\begin{pmatrix}\partial\beta\\ \partial\Omega^T\end{pmatrix}=\begin{pmatrix}\partial\beta\\ \partial\Omega\end{pmatrix}^T \quad (41)$$

$$\begin{pmatrix}\partial\Omega\\ \partial\kappa^T\end{pmatrix}_{ij}=\frac{\partial\Omega_i}{\partial\kappa_i} \quad (42)$$

The entry is in the ith row and jth column $\frac{\partial\Omega_i}{\partial\kappa_i}$. It is completely noticeable that it converts into a chain-like argument. κ to Ω. The chain rule computes:

$$\begin{pmatrix}\partial\beta\\ \partial\kappa^T\end{pmatrix}, as\begin{pmatrix}\partial\beta\\ \partial\kappa^T\end{pmatrix}=\begin{pmatrix}\partial\beta\\ \partial\Omega^T\end{pmatrix}\begin{pmatrix}\partial\Omega\\ \partial\kappa^T\end{pmatrix} \quad (43)$$

A loss function can be used to investigate the disparity between a CNN's predicted outcome and the real outcome κ^L and the target $t,\kappa^1\to w^1, t,\kappa^2\to\ldots..,$ $t,\kappa^L\to w^L=\beta,$,. The utilisation of complex functions is common. Prediction results are $\arg\max_i \kappa_i^l$. Convolution process:

$$\Omega i^{l+1}, j^{l+1}, d=\sum_{i=0}^{h}\sum_{j=0}^{w}\sum_{k=0}^{d}F_{i,j,k}\times\kappa_{i^{l+1}}^{L}+, ij^{l+1}+j,k \quad (44)$$

$\Omega\left(\kappa^{l+1}\right)_{\text{in}}$ $\mathrm{P}^{h^{l+1}\times w^{l+1}\times d^{l+1}}$,

$$h^{l+1}=h^l-h+1, w^{l+1}=w^l-w+1, d^{l+1}=d \quad (45)$$

Each training example's label in Inception V3 $\eta\in\{1,\ldots n\}$ probability is calculated as:

$$P(\eta\mid n)=\frac{\exp(z_k)}{\sum_k^i\exp(z_k)}, \quad (46)$$

where β shows an unnormalized log probability. Normalising $q(\eta\mid n)$ Truth distribution among labels $\sum_k q(\eta\mid n)=1$. This model's loss is given by cross-entropy..

$$l=\sum_{k=1}^{K}\log(P(\eta))q(\eta) \quad (47)$$

The following basic gradients are used by cross-entropy loss, a logit-independent technique, is used while training deep models:

$$\frac{\partial l}{\partial z_k}=P(\eta)-q(\eta) \quad (48)$$

The range of the following equation increases the log-likelihood of an accurate label while decreasing cross entropy. It is possible to analyze a label distribution without considering training occurrences using Inception V3 $u(\eta)$ With a working outsample and a smooth parameter. $q(\eta\mid n)=\vartheta_{\eta\mid,\Omega}$ is updated by:

$$q^{l(\eta\mid n)}=(1-\in)\vartheta_{\eta,n}+\in u(\eta), \quad (49)$$

It blends distributions with initial and set weights. A uniform distribution label-smoothing regularisation $u(\eta)=1/K$ yields:

$$q^{l(\eta\mid n)}=(1-\in)\vartheta_{\eta,n}+\frac{\in}{K}, \quad (50)$$

The following definition of cross-entropy.

$$h(q',P)=-\sum_{k=1}^{K}\log(P(\eta))q^{l(k)}=(1-\in)h(q',P)+\in h(u,P) \quad (51)$$

Using a single cross entropy loss is identical to label-smoothing regularisation $h(q,P)$ with a couple of losses $h(q,P)$ and $h(u,P)$. Loss two impairs label dissemination P deviation with weight from prior $\frac{\in}{\in+1}$, It complements to finding the Kullback-Leibler divergence. Now, based on those mentioned perceptions, the deep convolutional neural network and the TL model

from the source domain are combined, as illustrated in figure 3.

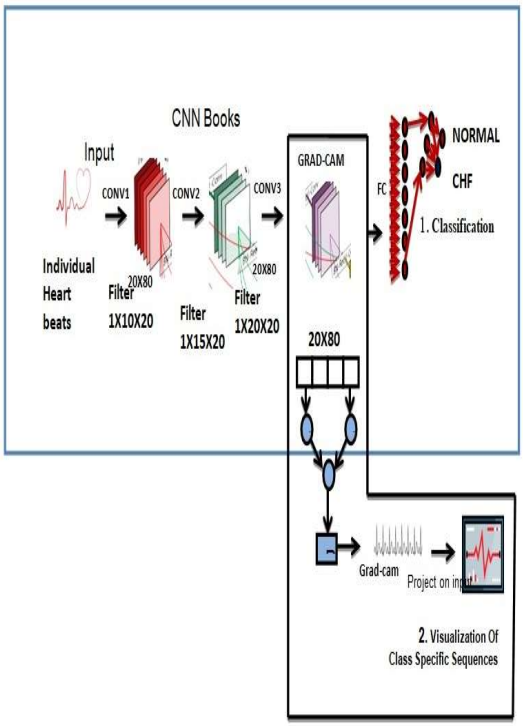


FIGURE 3. ACLS-RCNN ARCHITECTURE

3.5.1. The Improved Cuttle-fish Swarm Optimization Algorithm (ICSOA)

Cuttlefish are cephalopods that can alter their colour to blend in or perform. The three cell layers of cuttlefish, which alter their skin colour, are crucial to the most effective treatment. In cephalopods, various cell layers, colours, leucophores, and tints produce various tones. The cuttlefish's method of reflecting light is covered through reflection interaction. Here are some tips on how to use reflection and perceivability to spot a different layout (new:

$$newq=visibility+reflection \quad (52)$$

P (creatures) of N possible configurations. The conveying of the d-layered challenge area is random;

$$Q[k].points[s]=random*(upperLimit)+ lowerLimiti = 1, 2\ldots N; j = 1, 2\ldots d \quad (53)$$

in which any integer that falls between upperLimit and lowerLimit is considered irregular, which represent the problem space's higher and lower boundaries, respectively (2, 3).

Each location within the population is given consideration for viability and serves as a component of M's ground truth. The switching of photoreceptor and iridophores (light) units causes reflected fluctuation. Each chromatophore cell tightens or loosens its filaments, shortening or lengthening the model of impossible. The representations of these exercises come after those sections.

$$visibility=W*(Bestpoints[s]-G\_1[k].Points[s] \quad (54)$$

$$reflection\_s=A*I\_1\ [k].Points\ [s] \quad (55)$$

Performed the process described in the previous equation using cells. Cell I therefore belongs to the G1 team. The j-th component of the I-th cell is indicated by the symbol points[j]. Using Best Focuses, the optimal arrangement focuses are displayed. When the muscle of the cell protracts or relaxes, the stretch range of the scale is determined by the reflection degree (R). The exact perceivability level of the example is indicated by the letter V. To find the R and V properties, take the following actions:

$$W=random()*(w\_2-w\_2\ )+w\_2 \quad (56)$$

$$A=random()*(a\_(1-)\ a\_2\ )\ a\_2 \quad (57)$$

The constant integers v1 or v2 control the degree stretch in the final view. Using the r1 and r2 model bounds, the stretch time period is sorted. In certain circumstances, all that is necessary is setting of either V or R value to 1.

4. PERFROMANCE ANALYSIS

The review, f-score, and precision of our proposed study for coronary illness prediction are presented in this section. Our method just guesses the information; therefore we group them based on their expected results before using the information as excellent situations.

Journal of Theoretical and Applied Information Technology
15th May 2024. Vol.102. No 9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A. Accuracy

How closely a result computation, estimation or conclusion adheres to the factual value or standard. We may carefully scrutinise the instances arranged with their results, as suggested by exactness. Under Fig. 4, the accuracy of the traditional and proposed technique is shown.

$$Accuracy=((A + B))/((A + B + C + D)) \quad (58)$$

True positive is denoted by "A=amount of right figures of a positive example", true negative by "B=amount of right estimates of a negative example", phoney positive by "C=amount of wrong conjectures of a positive example", and false negative by "D=amount of wrong gauges of a negative example".

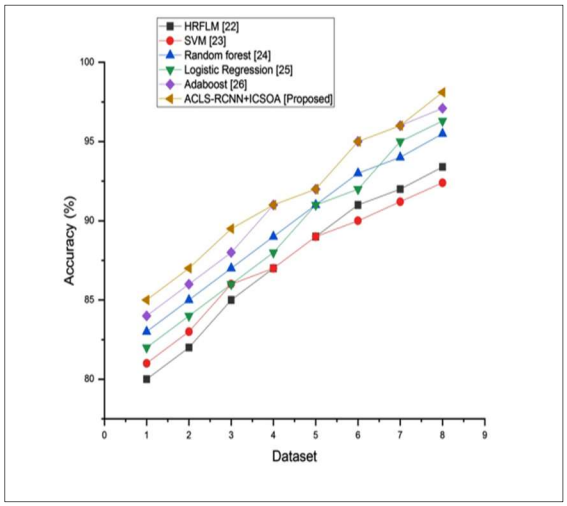


FIGURE 4. ACCURACY COMPARISON

B. Precision

Regardless of whether the estimates are true or not, this accuracy is referred to as "the property of being exact" and pertains to at least two estimates. In comparison to other typical outcomes, this determines how precise our results are. The precision of the suggested approach is shown in Figure 5.

$$precision=A/((A + C)) \quad (59)$$

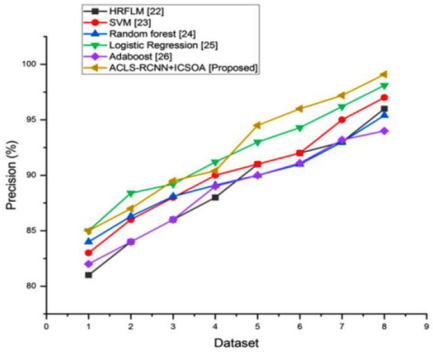


FIGURE 5. PRECISION COMPARISON

C. Recall

The degree of the total quantity of significant instances acquired is the review, also known as awareness. This boundary facilitates the identification of the trial-related instances obtained. The recall comparison is shown in Figure 6.

$$Recall=A/((A + D)) \quad (60)$$

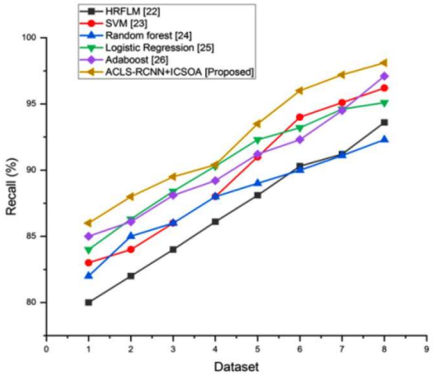


FIGURE 6. RECALL COMPARISON

D. F1 - score

This F1-score is the exactness and review symphony. Utilising measurements, things are rated. Figure 7 below displays the outcomes.

$$f1\text{-}score=A/(A + 0.5*(C+D)) \quad (61)$$

Journal of Theoretical and Applied Information Technology
15th May 2024. Vol.102. No 9

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

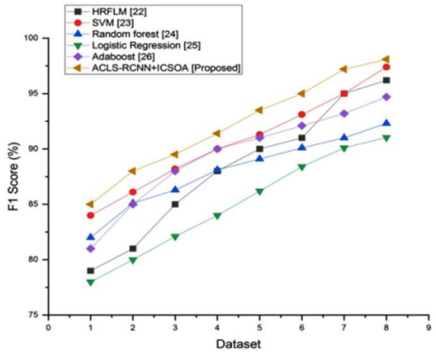


FIGURE 7. F1-SCORE COMPARISON

Discussion

Using the provided data, we examine the suggested approach by comparing exhibition metrics with coronary disease expectancy. The analysis shows that the "Kaggle information base" was used to put together the datasets. The provided datasets have been balanced and normalized using standardization techniques to eliminate any undesired anomalies and disruptions. The outcomes of this method are then used to build prepared and tested assortments. To predict cardiac disorders, the generated assortment is put through additional cycles like highlight determination, component extraction, and grouping. Along with being used, the suggested technique is coordinated with other industry-accepted practices, including HRFLM [22], SVM [23], Arbitrary Woods [24], Strategic Relapse [25], and AdaBoost [26]. The aforementioned measures can be arranged by calculations utilizing standard or recommended procedures.

The resulting intricacies draw attention to the weaknesses of current methods. Expectation techniques can be advanced by using AI computations in the following ways [22]. Furthermore, novel element determination methods may be developed to better understand the most crucial indicators of cardiac disease. By no means does the SVM method in [23] work incredibly well for large datasets. SVM has trouble applying itself when the dataset include extra noise, such as Cross-over. The SVM may become discouraged if the amount of attributes per information point significantly exceeds the preceding information tests. Slope helped trees forecast complex problems in [24], although their accuracy was frequently lower. A forest is more complex to comprehend than one selected tree. It is easy to practice, comprehend, and execute strategic relapse [25]. Generally speaking, when the quality totals of the informative collection exceed the totals of the informational focus, do not use strategic relapse. In terms of the element space section, no one might be anticipated. AdaBoost typically requires a large dataset [26]. Before utilizing an AdaBoost, it's imperative to remove noise and exceptions.

Based on this evaluation, we assume that the suggested system is better than the others. One minor change could have a big impact on the Irregular Backwoods. It is possible that the Irregular Woods method computations are much more complex than other calculations. Regarding methodology, however, they think that our approximations will be more useful for instant access and that information change is obviously not a major concern. If there are more highlights than exactly the number of information, then calculated relapse should not be employed; however, our proposed method was looked at. On the other hand, this could lead to mistakes. While strategic relapse creates boundaries, our proposed approach does not.

Table 1: Evaluation of the Proposed on NPV and MCC

Method	NPV	MCC
Proposed Method	94.6	92.8
HRFLM	91.7	88.4
SVM	90.3	86.6
Random Forest	86.9	86.8
Logistic Regression	84.4	84.0

Table 1 lists metrics such as MCC and the NPV detection model. By contrasting the proposed model with the existing methods, a higher metrics value was found. The recommended model's low error rate, dependability, and ability to handle confusing data are indicated by the more favorable MCC and NPV values.

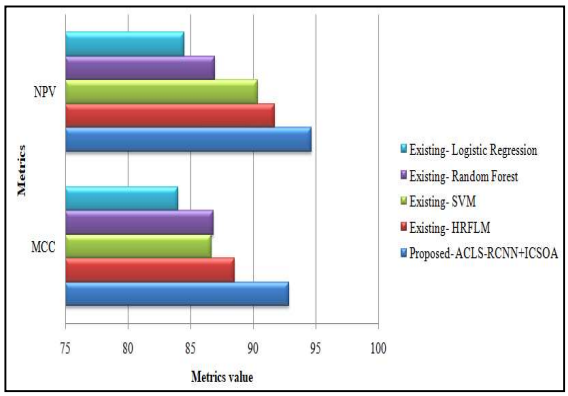


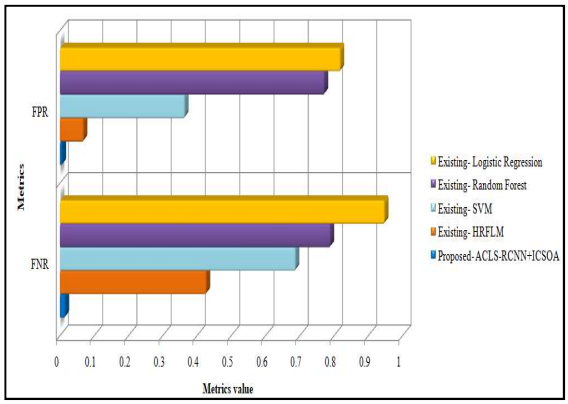
FIGURE 8. PROPOSED ACLS-RCNN+ICSOA DEMONSTRATION USING NPV AND MCC

The NPV, MCC, and f-measure metrics of the suggested ACLS-RCNN+ICSOA are shown in Figure 8. Comparing the measurements to HRFLM, SVM, Random Forest, and Logistic Regression is the next stage. With a design, NPV, MCC, and f-measure ought to increase. The method produces 94.63% NPV and 92.83% MCC. Metric values ranging from 86.94% to 93.67% for the current techniques were lower than those for the proposed methodology. The proposed ACLS-RCNN+ICSOA method predicts cardiac disease effectively and with high accuracy.

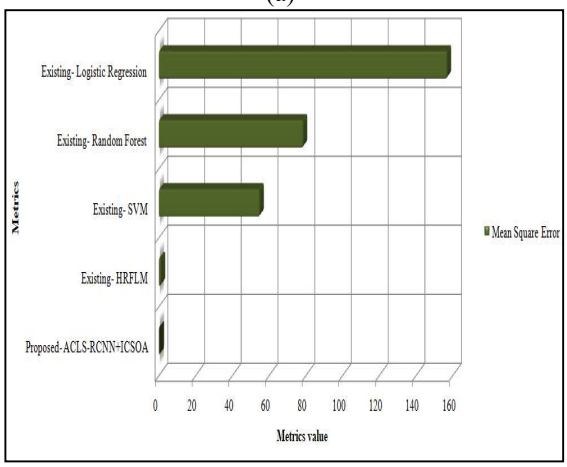
Table 2: Focused on FPR, FNR, and proposed method

Method	FNR	FPR	Computation Time	MSE
Proposed Method	0.014	0.005	20345	0.005
HRFLM	0.427	0.068	48381	0.979
SVM	0.687	0.364	66434	54.78
Random Forest	0.789	0.771	78946	78.45
Logistic Regression	0.947	0.818	89768	156.79

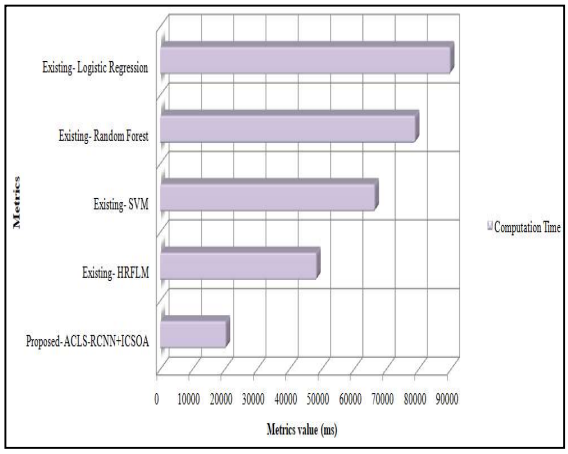
The prediction models for heart disease calculation time, FPR, MSE, and FNR are among the metrics shown in Table 2. Compared to previous strategies, the recommended strategy outperformed them. The proposed technique achieves low FNR and FPR between expected and actual class values. A low FNR value indicates an excellent model. The ratio of actually negative test findings to all negative test results—including data that was mistakenly deemed unrelated to traffic noise—is known as the mean square error, or MSE. This property may help determine the likelihood that a negative test result is noise.



(a)



(b)



(c)

FIGURE 9. DEMONSTRATION OF PROPOSED METHOD BASED ON FNR, FPR, MSE, AND COMPUTATION TIME

Figure 9 displays the suggested method's graphical analysis in terms of computation time, FNR, FPR, and MSE utilizing Random Forest, HRFLM, Logistic Regression, and SVM. MSE is used to illustrate the diversity of the dataset used

for heart disease and the dependability of the suggested prediction methods. The new study has a low MSE value, indicating a less effective design, compared to the previous technique's high MSE value. We analyze the proposed technique and evaluate the misclassification risk using FPR and FNR measures. The proposed strategy results in lower FNR and FPR measurements compared to the present methodologies, leading to misprediction. Consequently, the proposed method avoids misprediction of heart disease and has a short calculation time.

5. CONCLUSION

This article describes an intelligent ACLS-RCNN-based cardiac disease prediction system that makes use of optimal structures to increase accuracy. In the first step of the hybrid optimization process CI-AO+GI-SMO, multivariate statistical analysis is used to choose the most significant features from the original population. The search method combines an AO for local search with a modified SMO for global search. AO is also put into practice the concept of rehabilitation for candidates who were rejected. The area under the ROC curve of the proposed approach is evaluated together with its specificity, sensitivity, and accuracy with use of publicly accessible datasets. The experiment's outcomes showed that the ACLS-RCNN-based ICSOA technique approach had high prediction accuracy for heart sickness on the four datasets, ranging from 95.6% to 96.4%. Additionally, the suggested approach outperformed the ultra modern prediction techniques. The recommended technique is examined at the image level for a more detailed examination. A thorough and reliable diagnosis of cardiac disease is made possible by the research. The results indicate that our proposed technique will support in the early identification of heart issues.

REFERENCES:

[1] Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J.T., 2018. Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics, 19(6), pp.1236-1246.

[2] R. Das, I. Turkoglu, and A. Sengur, ‘‘Effective diagnosis of heart disease through neural networks ensembles,’’ Expert Syst. Appl., vol. 36, no. 4, pp. 7675–7680, 2009.

[3] Buja LM, McAllister Jr HA. Coronary artery disease: pathological anatomy and pathogenesis. In: Willerson JT, Cohn JN, Wellens HJJ, Holmes Jr DR, editors. Cardiovascular medicine, 3rd ed. London: Springer, pp. 593-610, 2007.

[4] Z. Arabasadi, R. Alizadehsani, M. Roshanzamir, H. Moosaei, and A. A. Yarifard, ‘‘Computer aided decision making for heart disease detection using hybrid neural network-genetic algorithm,’’ Comput. Methods Programs Biomed., vol. 141, pp. 19–26, Apr. 2017.

[5] H. Yan, Y. Jiang, J. Zheng, C. Peng, and Q. Li, ‘‘A multilayer perceptron based medical decision support system for heart disease diagnosis,’’ Expert Syst. Appl., vol. 30, no. 2, pp. 272–281, 2006.

[6] K. Vanisree and J. Singaraju, ‘‘Decision support system for congenital heart disease diagnosis based on signs and symptoms using neural networks,’’ Int. J. Comput. Appl., vol. 19, no. 6, pp. 6–12, 2011

[7] Ali, L., Rahman, A., Khan, A., Zhou, M., Javeed, A. and Khan, J.A., 2019. An automated diagnostic system for heart disease prediction based on ${\chi^{2}} $ statistical model and optimally configured deep neural network. IEEE Access, 7, pp.34938-34945.

[8] Gupta, A., Kumar, R., Arora, H.S. and Raman, B., 2019. MIFH: A machine intelligence framework for heart disease diagnosis. IEEE Access, 8, pp.14659-14674.

[9] Paul, A.K., Shill, P.C., Rabin, M., Islam, R. and Murase, K., 2018. Adaptive weighted fuzzy rule-based system for the risk level assessment of heart disease. Applied Intelligence, 48(7), pp.1739-1756.

[10] Tuli, S., Basumatary, N., Gill, S.S., Kahani, M., Arya, R.C., Wander, G.S. and Buyya, R., 2020. Health Fog: An ensemble deep learning based Smart Healthcare System for Automatic Diagnosis of Heart Diseases in integrated IoT and fog computing environments. Future Generation Computer Systems, 104, pp.187-200.

[11] Ali, F., El-Sappagh, S., Islam, S.R., Kwak, D., Ali, A., Imran, M. and Kwak, K.S., 2020. A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. Information Fusion, 63, pp.208-222.

[12] Acharya, U.R., Fujita, H., Oh, S.L., Hagiwara, Y., Tan, J.H., Adam, M. and Tan, R.S., 2019. Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals. Applied Intelligence, 49(1), pp.16-27.

[13] Chowdhury, M.E., Khandakar, A., Alzoubi, K., Mansoor, S., M Tahir, A., Reaz, M.B.I. and Al-Emadi, N., 2019. Real-time smart-digital stethoscope system for heart diseases monitoring. Sensors, 19(12), p.2781.

[14] Ketu, S. and Mishra, P.K., 2022. Empirical analysis of machine learning algorithms on imbalance electrocardiogram based arrhythmia dataset for heart disease detection. Arabian Journal for Science and Engineering, 47(2), pp.1447-1469.

[15] Jalali, S.M.J., Karimi, M., Khosravi, A. and Nahavandi, S., 2019, October. An efficient neuroevolution approach for heart disease detection. In 2019 IEEE international conference on Systems, Man and Cybernetics (SMC) (pp. 3771-3776). IEEE.

[16] Atallah, R. and Al-Mousa, A., 2019, October. Heart disease detection using machine learning majority voting ensemble method. In 2019 2nd international conference on new trends in computing sciences (ICTCS) (pp. 1-6). IEEE.

[17] Deperlioglu, O., Kose, U., Gupta, D., Khanna, A. and Sangaiah, A.K., 2020. Diagnosis of heart diseases by a secure Internet of Health Things system based on Auto encoder Deep Neural Network. Computer Communications, 162, pp.31-50.

[18] Olaniyi, E.O., Oyedotun, O.K. and Adnan, K., 2015. Heart Diseases Diagnosis Using Neural Networks Arbitration. International Journal of Intelligent Systems & Applications, 7(12)

[19] Waqar, M., Dawood, H., Dawood, H., Majeed, N., Banjar, A. and Alharbey, R., 2021. An efficient smote-based deep learning model for heart attack prediction. Scientific Programming, 2021.

[20] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, *7*, 81542-81554.

[21] Fitriyani, Norma Latif, Muhammad Syafrudin, GanjarAlfian, and Jongtae Rhee. "HDPM: an effective heart disease prediction model for a clinical decision support system." *IEEE Access* 8 (2020): 133034-133050.

[22] Khan, Mohammad Ayoub. "An IoT framework for heart disease prediction based on MDCNN classifier." *IEEE Access* 8 (2020): 34717-34727.

[23] Gárate-Escamila, A. K., El Hassani, A. H., &Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, *19*, 100330.

[24] Ali, Liaqat, Atiqur Rahman, Aurangzeb Khan, Mingyi Zhou, Ashir Javeed, and Javed Ali Khan. "An automated diagnostic system for heart disease prediction based on ${\chi^{2}} $ statistical model and optimally configured deep neural network." *IEEE Access* 7 (2019): 34938-34945.

[25] Dutta, Aniruddha, TamalBatabyal, MeheliBasu, and Scott T. Acton. "An efficient convolutional neural network for coronary heart disease prediction." *Expert Systems with Applications* 159 (2020): 113408.

[26] K. L. Anusha, Janardhan G, Mahankali Saritha, and Manne Archana. "Artificial Neural Network with Spider Monkey Optimization Algorithm for Cardiovascular Disease Prediction." Journal of Theoretical and Applied Information Technology, 101(17) 2023. pp. 7096-7105.

[27] Krishna Lava Kumar Gopu, and Suthendran Kannan. "Deep Graph Neural Network with Fish-Inspired Task Allocation Algorithm for Heart Disease Diagnosis". Journal of Current Science and Technology, Vol. 13, Issue 2, 2023. pp. 392-411.