FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

Hybrid Optimization based Feature Selection with DenseNet Model for Heart Disease Prediction

Dr. V. Gokula Krishnan[1*], Dr. M. V. Vijaya Saradhi[2], Dr. S. Sai Kumar[3], G. Dhanalakshmi[4], P. Pushpa[5] and Dr. V. Vijayaraja [6]

[1] *Professor, Department of CSE, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Thandalam, Chennai, Tamil Nadu, India, gokul_kris143@yahoo.com*
[2]*Professor & Head, Department of CSE, ACE Engineering College, Ghatkesar, Hyderabad, Telangana, India, meduri.vsd@gmail.com*
[3]*Assistant Professor, Department of IT, PVP Siddhartha Institute of Technology, Vijayawada, Andhra Pradesh, India, saikumar@pvpsit.ac.in*
[4]*Associate Professor, Department of IT, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India, dhanalakshmi4481@gmail.com*
[5]*Assistant Professor, Department of AIDS, Rajalakshmi Institute of Technology, Kuthambakkam, Chennai, Tamil Nadu, India, ppushpacse88@gmail.com*
[6]*Department of AIDS, R M K College of Engineering and Technology, Kavaraipettai, Tamil Nadu, India, vijayarajaads@rmkcet.ac.in*

*__Correspondence:__ Dr. V. Gokula Krishnan; gokul_kris143@yahoo.com

ABSTRACT- The prevalence of cardiovascular diseases (CVD) makes it one of the leading reasons of death worldwide. Reduced mortality rates may result from early detection of CVDs and their potential prevention or amelioration. Machine learning models are a promising method for identifying risk variables. In order to make accurate predictions about cardiovascular illness, we would like to develop a model that makes use of transfer learning. Our proposed model relies on accurate training data, which was generated by careful Data Collecting, Data Pre-processing, and Data Transformation procedures. Additionally, the optimal selection is carried out on the existing attributes by fusing two meta-heuristic procedures, the Lion Algorithm (LA) and the Butterfly Optimization Procedure (BOA), into a single method dubbed the hybrid Lion-based BOA (L-BOA). In this research, we analyse the amount of parameters in a deep learning model and provide an end-to-end solution for classifying patients as healthy or unwell. To extract deep features from the best data, the proposed approach makes use of pre-trained convolutional neural networks-(CNNs) dubbed DenseNet121. The model can benefit from a more nuanced feature set composed of the features derived from each CNN. Accuracy, precision, recall, and the F1-Score were used to rate the trained classifiers. The models' classification results demonstrated that the inclusion of pertinent characteristics significantly improved the classification precision. When compared to models skilled on a full feature set, the performance of organization replicas trained with a smaller feature set improved dramatically with less training time.

Keywords: Cardiovascular diseases; Data Pre-processing; Convolutional neural networks; Butterfly Optimization Algorithm; Early diagnosis.

ARTICLE INFORMATION

Author(s): Dr. V. Gokula Krishnan, Dr. M. V. Vijaya Saradhi, Dr. S. Sai Kumar, G. Dhanalakshmi, P. Pushpa and Dr. V. Vijayaraja;

Received: 31/03/2023; **Accepted**: 13/04/2023; **Published**: 30/04/2023
e-ISSN: 2347-470X;
Paper Id: IJEER 3103-07;
Citation: 10.37391/IJEER.110203
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-11/ijeer-110203.html

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Traditional and labour-intensive techniques of storing and analysing medical data are all that hospitals and clinics have at their disposal. In an effort to overcome this barrier, many medical institutions have combined vast data resources with cutting-edge technologies [1]. Despite these efforts, many hospitals and clinics have been slow to adopt the new systems. Despite the vast amounts of information, there is still a lot we don't know about illnesses and how to treat them. Because of the complexities of the data, (ML) approaches [2] are gaining ground in the field of data investigation. The combination of ML and data-driven strategies can yield reliable diagnostic instruments. The purpose of this research is to identify and evaluate the institutional barriers that prevent medical organizations from implementing a proven strategy that provides management with strategic answers to these challenges [3]. Predicting the start of cardiac disease is one of the most pressing and difficult problems facing modern medicine. In many industrialised countries, heart disease is the number one killer. Around 3% of the overall health care budget is used to treat heart disease [4], and yet half of all patients diagnosed with 1-2 years. Several tests are needed for heart disease prediction. Early diagnosis can be challenging, and

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

erroneous predictions made by medical staff due to a lack of experience can increase patient anxiety. Because to a shortage of educated equipment, and other resources necessary for efficient diagnosis and management of patients difficult in developing states [5]. Preventing life-threatening heart attacks and increasing patient safety depend on an accurate assessment of failure. With the right training data, ML algorithms can perform well in disease diagnosis. There are freely available datasets for cardiovascular disease prediction model comparison [6]. Researchers are able to make better use of the available massive databases thanks to the advent of ML and artificial intelligence.

The importance of lowering mortality from cardiovascular diseases has been underscored by recent studies focusing on heart glitches in adults and children. Given the inconsistency and redundancy of the available clinical datasets, pre-processing is essential. It is crucial to choose the relevant features that can serve as risk variables in forecast models. To create reliable prediction models, it is significant to carefully choose the optimal combination of features and ML procedures [7]. Evaluating the impact of risk variables that match the three criteria of high prevalence in maximum populations, independent impact on heart disease risk, and controllability or treatability is crucial [8]. Many risk factors and clinical characteristics have been included into CVD prediction models. Features used in the development of CVD risk assessments family history, and heart status.

According to recent research [11], at least 14 characteristics are required for accurate and reliable prediction. Researchers nowadays are struggling to identify the right mixture of characteristics and machine learning methods to reliably predict cardiovascular disease. Training machine learning algorithms on high-quality data is essential for optimal performance [12, 13]. Feature selection methods prepare the data in order to offer a more accurate forecast [14], as the procedures depend on the consistency of the training and test data. Classifiers and hybrid models can then be used to provide predictions about the likelihood of illness occurrence after the relevant features have been picked. Researchers have used a wide variety of methods to create classifiers and mixture models [15, 16]. Limited medical datasets, feature selection, ML procedure applications, and a lack of in-depth research are just few of the problems that may prohibit effective prediction of heart disease. Our study's goal is to fill up approximately of these knowledge gaps so that we can create a more accurate model for forecasting CVD.

Predicting when someone may develop a disease is a difficult problem that needs to be solved before the disease reaches a critical stage. Individuals are safer after the sickness has been predicted. The cost-effectiveness of early stage prediction is not to be underestimated. The hybrid L-BOA algorithm is used to achieve optimal feature selection, which reduces the time it takes to high-level landscapes from the input data. The use of a pre-trained model of CNN allows for extremely precise predictions to be made. Below is the outline for the rest of the paper: *Section 2* covers the cited works, while *Section 3* details the proposed model. In *section 4*, the validation analysis is presented, and in *Section 5*, the findings are summarised.

2. RELATED WORKS

To aid in the detection of hypertension due to inherited heart disease, Ge et al. [17] created a computer aided analysis based on a single cycle with several features. This easy-to-use, non-invasive technique has the potential to aid in the detection of CHD-PAH at an early period. To begin, we pre-processed the raw heart sounds by separating out individual cardiac cycles using an adaptive double-threshold segmentation technique. The characteristics of the cardiac cycle and the S2 component are then retrieved using the domain and the wavelet packet energy domain. In addition, a (CNN) is employed to extract the cardiac cycle's depth information. A fused feature vector was created by combining the aforementioned features. XGBoost was used as the classifier to separate samples into the normal, CHD, and CHD-PAH categories. To get the best categorization result for many outcomes representing multiple cardiac cycles of the same person, the majority voting technique is utilised. With this fresh strategy, we were able to improve classification accuracy to 88.61 percent.

To improve the accuracy of heart disease analysis categorization, Pradhan and Bhuiyan [18] offer an adaptive bacterial foraging optimisation (ABFO) procedure for feature selection. The cardiac image is first processed through a median filter, which eliminates any distracting background noise. The cardiac image is then processed to isolate GLCM characteristics. Using the ABFO method, the best features are designated from the set of extracted features. The classifier, a support vector neural network, receives the selected features as input (RBNN). The image is sorted by the classifier into normal and pathological categories. Simulation consequences demonstrate that the ABFO-based RBNN outperforms its predecessors—the standard RBNN, the ANN, the KNN, and the SVM.

(ANNs) have grown in popularity thanks to their accurate predictions, and Hassan et al. [19] offers one such network. So, it's important to study ways to close this gap so that better forecasts can be made. In light of this, this study suggests a novel method for HD prediction, one that makes use for feature extraction, Principal Component for prediction. The suggested method (DNN + PCA + LR) was tested on Cleveland, a freely available HD dataset. The suggested method achieves high levels of accuracy (91.79%) on training data and (93.33%) on testing data, as shown by experimental findings. In addition, the proposed method outperformed the state-of-the-art methods across the board in rapports of evaluation criteria.

Based on machine learning tactics, Nagarajan et al. [20] suggested a system for the rapid and precise diagnosis of cardiovascular illness. As heart disease was identified as a major health risk, it is essential that screenings and diagnoses be performed remotely and routinely to allow for timely intervention. Several models have been presented in the past year to attempt to estimate how common heart disease is in the modern world. The optimisation algorithm is critical for precise diagnosis of cardiac disease. Using deep convolutional neural networks, this study aimed to create a hybrid GCSA that would reflect a genetic-based crow search strategy for classification.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

Results demonstrate that the suggested model GCSA improves classification accuracy over the other feature selection approaches by more than 4%.

Khan et al. [21] propose a method for representing PCG signals as non-stationary that makes use of the Fourier-Bessel series development of the wavelet transform (FBSE-EWT). Finally, we may use the 2D-PSRs to isolate the most crucial Fourier-Bessel functions (FBIMFs) and learn about their underlying geometric features. We were able to lessen the burden by shrinking the feature set with the help of metaheuristics optimization-based features selection (FS) algorithms like the salp swarm optimisation algorithm (SSOA), the emperor penguin optimisation procedure (EPOA), and the tree growth optimisation algorithm (TGO) (TGOA). Several FS methods have been analysed, and their results associated to those of machine learning classifiers. This demonstrates that PCG signals may be effectively categorised using the FS techniques used to extract features using FBSE-EWT. With the reduced features set created by SSOA, the projected method has obtained the highest classification accuracies possible, with values as high as 99.70% for the two-class problem and as high as 99.70% for the five-class problem. So, not only did this technology benefit cardiologists, but it also aided in the creation of wearable cardiac devices for monitoring and diagnosing heart conditions (since it employs a reduced feature set).

Manimurugan *et al.* [22] suggested a two-stage system for classifying and forecasting medical data. If the First Stage outcome was sufficient for forecasting cardiovascular disease, there would have been no need for Stage 2. Classification of echocardiogram pictures for heart disease prediction followed cataloguing of data acquired by medical sensors put on the patient's body in a two-stage process. Echocardiography picture classification was completed using a hybrid of the Faster R-CNN and the SE-ResNet-101 model, while classification of sensor data was achieved using a hybrid of linear discriminant examination and the modified ant lion optimisation (HLDA-MALO). Results from two separate categorization methods were integrated and checked for their accuracy in identifying at-risk individuals for cardiovascular disease. Normal sensor data was identified with 96.85% accuracy using the HLDA-MALO method, while aberrant data was identified with 98.31% correctness. The recommended hybrid Faster R-CNN with SE-ResNeXt-101 transmission learning model achieved the highest accuracy (99.15%) when used to the classification of images, with outcomes of 98.06% precision, 98.95% recall, 96.32% 99.02% F-score.

3. PROPOSED METHODOLOGY

This study demonstrates how crucial feature selection is for precise cardiac illness classification. *Figure 1* provides the workflow of the proposed model.

3.1 Datasets

The purpose of this work was to construct a (ML)-based system for detecting heart disease by analysing the impact various variables have on the incidence of heart disease, using two datasets known as Framingham. Two datasets are utilised to investigate a wide range of diagnostic characteristics and clinical procedures for stroke identification in the human body.

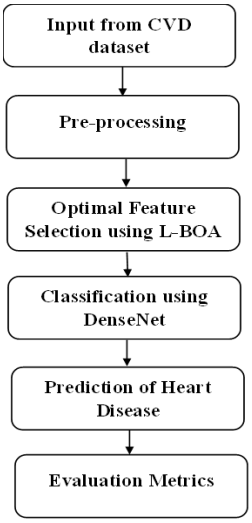

Figure 1: Workflow of the Proposed Model

The data sets came from many places. Age, other key medical parameters were included in the datasets, all of which are directly associated to the incidence of illness and offer a great deal of flexibility for analysing heart disease. There were two considerations in selecting the datasets. First, we needed a criterion that would allow us to compare and contrast the many medical techniques available for treating cardiac disease. Second, we choose our datasets with an eye on data completeness and accessibility. The quantity and variety of data and attributes present in datasets can vary widely depending on their origin. So, we choose datasets that provided a sufficient amount of data and showed some degree of feature parity with one another

3.1.1 CVD

As part of its healthcare hackathon [23], McKinsey & Company is in charge of the CVD dataset. A free dataset repository [24] is available where the dataset can be downloaded. There were 29072 patient observations with 12 data topographies in the gathered dataset. Eleven of them are typical clinical symptoms and are therefore input characteristics, while the 12th, "stroke," is the target feature that indicates whether or not the patient has had a stroke. *Table 1* provides a comprehensive breakdown of the CVD dataset's data elements.

Table 1: Explanation of features CVD dataset

Description	Attribute
patient's i.d	i.d
includes (''male'': 0, ''female'': 1, ''other'': 2)	gender
patient's age	age
suffering from hypertension (''yes'':1, ''no'':0)	hypertension
suffering heart disease (''yes'':1, ''no'':0)	heart _disease
marital position of patient (''yes'':1, ''no'':0)	ever_married

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

job status (‘‘children’’:0, ‘‘govt_job’’:1, ‘‘never_worked’’:2, ‘‘private’’:3, ‘‘self_employed’’:4)	work_type
(‘‘rural:0, ’urban":1)	residence_type
regular glucose level of blood	avg_glucose_level
body mass index	bmi
(‘‘never smoked’’:0, ‘‘formerly smoked’’:1, ‘‘smokes’’:2)	smoking_status
(‘‘yes’’:1, ‘‘no’’:0)	stroke

3.1.2 Framingham

The Framingham dataset [25] is a collection of medical records from a cardiovascular study being conducted on the people of Framingham, Massachusetts. Patients' likelihood of developing CHD during the next decade is predicted using this dataset, which is primarily utilised in classification tasks. There are 4,240 patient records in all, with 15 attributes representing various potential dangers. The CHD was identified using 14 input factors. The data characteristics of the Framingham dataset are described in *Table 2*.

Table 2: Explanation of structures Framingham dataset

Description	Attribute
patient’s stage	age
(‘‘male’’:0, ‘‘female’’:1)	male
close of education (1 to 4)	education
(‘‘smoker’’:1, ‘‘non smoke’’:0)	Current Smoker
regular sum of cigarettes consumed per day	CigsPerDay
medicine (‘‘yes’’:1, ‘‘no’’:0)	BPMeds
previous stroke past (‘‘yes’’: 1, ‘‘no’’0)	prevalentStroke
hypertensive (‘‘yes’’:1, ‘‘no’’0)	prevalenHyp
previous diabetes history(‘‘yes’’:1, ‘‘no’0)	diabetes
cholesterol level	totCHol
systolic blood heaviness	sysBP
diastolic blood heaviness	diaBP
body mass index	BMI
heart rate measure	HeartRate
glucose level	glucose
target (‘‘yes’’: 1, ‘‘no’’: 0)	TenYearCHD

3.2 Pre-processing

One of the most crucial ladders in the machine learning procedure, data pre-processing facilitates subsequent analysis, and boosts the efficacy and efficiency of ML algorithms. Due to issues with imbalance in the obtained dataset, we performed certain pre-processing procedures. The CVD dataset has a total of 434,000 patient records out of which 14,754 had missing or null values. While in the Framingham dataset, out of 4240 patient records, 645 variables were missing. A null value could indicate that the value actually exist but cannot be determined at this time. Null or missing values in medical datasets are typically the result of incomplete data collection or the clinician's decision to disregard an observation because of the test's perceived low clinical use.

When dealing with missing data, data imputation techniques can be helpful; nevertheless, their application in medicine is restricted, and their specific efficacy for illness identification is unclear [26]. Since conventional data imputation techniques are unable to capture the missing data complexities in healthcare applications [27, 28], researchers often ignore the observations with missing values and knowingly discard the imperfect cases. Only a thorough understanding of a certain disease, however, can help choose an appropriate data imputation approach. In order to prevent biases in our results, we removed all the observations from both datasets that had a null value as per the above specified methodology.

Looking into the class distribution further, we see that both datasets were significantly imbalanced. Of the 29,072 patients in the CVD dataset, only 548 had stroke circumstances, while 28,524 did not. Only 557 out of 3101 patient records in the Framingham dataset had coronary heart disease risk information. As a result of the imbalanced data, classification mistakes occur when ML models are being trained [29]. Thus, we utilised a 'Random Down-Sampling' method to lessen the impact of imbalanced data. We divided the population into a minority group and a majority group. Patients who were experiencing heart problems were classified as a minority group, while those who were symptom-free were classified as the majority. There were 548 out of 28,524 total observations in the CVD dataset that were classified as minority. Using a sampling technique, we randomly selected 548 cases from the majority class and 548 examples from the marginal class, for a total of 1096 explanations in the dataset. For the Framingham dataset, the same procedure was used to produce 557 randomly selected observations from 3101 mainstream cases, for a grand total of 1114 explanations in a normal distribution. Effective feature importance and illness classification research necessitated the creation of two matched datasets.

3.3 Optimal Feature Selection

The suggested L-BOA method is applied on the feature attributes of the primary datasets. The total sum of features is given by the notation Fs n, where n is a positive integer between 1 and N. Optimal feature selection is indicated by the notation Fs (n*) opt, where n=1,2,...,N and N* is the total sum of optimum feature selections

3.3.1 Conventional Butterfly Optimization Procedure

BOA [30] acts similarly like butterflies when it comes to foraging and mating. These butterflies have been around for a long time because of their keen senses. They use all of their senses to find a mate and sustenance, including sight, touch, smell, hearing, and taste. These senses are crucial for getting from one area to another. Unique among the senses, smell aids the butterfly in its search for nourishment, usually nectar, even at great distances. The preparation of *fr* and the variation of A form the backbone of the butterfly process. Here, *fr* stands for scent, while A denotes the magnitude of the stimulus. The butterfly's encoded objective function is correlated with the intensity of the real-world stimulus A. Yet, the remaining butterflies can tell that the fragrance *fr* is relative. In this BOA algorithm, the scent is characterised in terms of the stimulus strength, as shown in *eq. (1).*

$$fr = smA^{pe} \quad (1)$$

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

In *eq. (1)*, *sm* stands for "sensory modality," and pe stands for "power exponent based on modality," both of which describe the varied degrees of absorption. In the classic BOA technique, the values for *sm* and *pe* are typically between 0 and 1. What follows is an idealised list of butterfly traits.

- The ability for butterflies to attract one another depends on a shared scent that they all emit.
- Each butterfly will fly to a different butterfly at random or towards the best butterfly in order to release its scent.
- The butterfly's response to stimuli is modified by the objective function.

Traditional BOA consists of three phases: "Initialization," "Iteration," and "Final." Each iteration begins with the start-up phase, then searches, and finally completes the process once an optimal solution has been found. Both the functions are established at the outset. Parameters used in this context are also assigned. Construct the basic butterfly population and optimise it later. The total number of butterflies and the amount of storage space allocated for simulation data are both fixed. The butterfly's positions, scent strength, and fitness are all created at random and stored in the search space.

The standard BOA algorithm is used to count the number of iterations during the iteration phase. Both the fitness of each butterfly and the number of times it can fly to a new location is calculated at each iteration. The algorithm initially determines the fitness values of all the butterflies at random positions in the solution space. In due time, these butterflies will generate the scent in their respective sites *via eq. (1)*. The model consists of two main phases, a local search phase and a global search phase. The corresponding equation, *eq. (2)*, depicts the butterfly's flight along the optimal solution path p* (2).

$$Li_i^{it+1} = Li_i^{it} + \left(rd^2 \times p^* - Li_i^{it}\right) \times fr_i \tag{2}$$

To solve *eq. (2)*, we plug in the solution vector for the i^{th} butterfly in round it as *Li* iit, the best solution from the current set of solutions as p*, a random sum between 0 and 1 as *rd*, and the fragrance of the i^{th} butterfly as *fr* i. The local search phase is represented by the equation *eq. (3)* in mathematics terms. In this case, the j^{th} and kth butterflies in repetition it are represented by *Li* jit and *Li* kit, respectively, from the solution space.

$$Li_i^{it+1} = Li_i^{it} + \left(rd^2 Li_j^{it} - Li_m^{it}\right) \times fr_i \tag{3}$$

Butterflies seek out a mate and feed in a wide variety of locations around the world. BOA utilises the switch probability *sp* to transition from a standard global search to a detailed local search.

3.3.2 Conventional Lion Procedure

The proposed L-BOA procedure is used for optimising hidden neurons in both NN and DBN. In the wild, a lion's primary drive is to establish and defend his territory, and this is something that LA [31] learns from its natural environment. Each of the six processing steps that make up the classic LA algorithm is briefly discussed below.

Generating with Pride: The migratory lion is designated by the initials "Li" and "Nom" whereas the territorial male lion and lioness are referred to by "Li" and "Ma" respectively. This wandering lion isn't related to the pride's current generation. The nomadic lions are the arbitrary numbers among the boundaries if a>1 and their vector rudiments are given by *li* ElMa, *li* ElFm, and the term li ElNorm signifies the solution vector representation of the lions. The length of the lion, Lng, is written as an equal sign (*Eq*) in the *equation (4)*.

$$Lng \begin{cases} a; & if\ a > 1 \\ b; & otherwise \end{cases} \tag{4}$$

The terms a and b in the preceding equation stand in for the integers required to determine the lions' length. When *a=1*, the procedure is forced to look for the lion in its binary representation, thus the members in the vector are either 1 or 0. Otherwise, it is evaluated using *eq. (5)*.

$$vlcty(li_{El}) \in (li_{El}^{min}, li_{El}^{max}) \tag{5}$$

After the pride has been generated, the next stage is to do a fertility evaluation. Here, the lion plays the role of the laggard, and the laggardness rate (denoted by the notation Lag Rt) is increased by one. Reference fitness is denoted by ftref if ft(LiMa)a is increasing. The territorial defence has occurred when the laggardness rate has exceeded Lag Rt max. Furthermore, the lioness' fertility is established by her sterility rate, str Rt, and is increased by 1 when str Rt is greater than str Rt's maximum value. When this occurs, the lioness is revised using *equations (6), (7), and (8)*.

$$li_{El}^{Fe} = \begin{cases} li_{fl}^{Fe+} & if\ El = fl \\ li_{El}^{FE} & otherwise \end{cases} \tag{6}$$

$$li_{fl}^{Fe+} = min\left[li_{fl}^{max}, max\left(li_{fl}^{min}, \Delta_{fl}\right)\right] \tag{7}$$

$$\Delta_{fl} = \left[li_{fl}^{Fe} + (0.1 rnd_2 - 0.05)\left(li_{fl}^{Ma} - rnd_1 li_{fl}^{Fe}\right)\right] \tag{8}$$

The following equations denote the random number between 1 and Lng by *fl*, the female updating function by, and random numbers between 0 and 1 by *rnd1* and *rnd2*, respectively. When the genetically modified lioness *Li* (Fe+) is mated with a normal male, the offspring are healthier. Until the maximum sum of female generations, fgn comax, is reached, the updating procedure continues. Once the upgrade is complete and Li (Fe+) is no longer available as a replacement for LiFe, it is checked to make sure that LiFe is the fertile one, so that the best cubs can be produced.

There are two main stages of mating and one additional stage. Here, mutation and crossing serve as the key processes, while gender clustering serves as an additional one.

The algorithm can avoid a local optimum and locate discrete solutions with identical fitness values by employing this straightforward approach for scanning the solution space, known as "territorial defence." When the criteria in *eq. (9), (10), and (11)* are met, the nomadic lion LiNom is selected.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

$$ft(Li^{Nom}) < ft(Li^{Ma}) \quad (9)$$

$$ft(Li^{Nom}) < ft(Li^{Mcb}) \quad (10)$$

$$ft(Li^{Nom}) < ft(Li^{fcb}) \quad (11)$$

When any of the two circumstances in *eq. (12) or Eq. (13)* is met, the algorithm is complete.

$$it > it_{max} \quad (12)$$

$$|ft(Li^{Ma}) - ft(Li^{Opt})| \leq er_{thrs} \quad (13)$$

In the aforementioned equations, the generation count is signified by it, which starts at 0, and increases by 1 upon achieving territorial dominance. The error tolerance level is denoted by the symbol er_{thrs}

3.3.3 Proposed L-BOA

Previous studies have combined various optimisation methods to create a novel hybrid optimisation algorithm. For some search issues, these provide the best possible outcomes. The benefits of rapid convergence are taken into account by this algorithm. The hybrid optimisation algorithm has demonstrated superior convergence behaviour when likened to other meta-heuristic algorithms. The standard BOA method can efficiently find solutions to the issues. Unfortunately, it has a number of drawbacks, including poor performance, early convergence, and a propensity to get stuck in local optima. The L-BOA procedure incorporates the LA algorithm into the BOA algorithm to address the shortcomings of BOA. The standard LA algorithm scales effectively and works well with high-dimensional problems; it also has a higher probability of discovery the global or adequate global optimum. The standard BOA algorithm updates using *eq. (2)* if rdsp and *eq. (3)* otherwise (3). The suggested L-BOA reverts to the standard BOA updating procedure if the chance sum rd is less than the switch likelihood sp. The female lion would then use *eq.* to execute the updating process (8).

3.4 Classification Process

The extracted feature maps are then utilised to discover an association between those characteristics and the desired output classes, which typically involve fully connected layers. When a layer is fully connected, all of its inputs are wired to each activation node in the layer above it. They are implemented as a classifier at the end of the network, in a layer with a number of neurons equal to the sum of output classes. The SoftMax function is used to convert the odds that an input belongs to a given class into decimal values in multi-class classification problems. The most likely class is chosen as the inputs' target anticipated group. The mathematical version of the SoftMax activation function for a K-class classification task is (14).

$$\sigma(Z)_i = \frac{e^{Zi}}{\sum_{i=1}^{k} e^{Zi}} \quad (14)$$

x *i* is the output along the i[th] dimension, I is a class sum from 1 to K, and (Z) *i* is the chance that the input is a member of *i*.

3.4.1 Transfer Learning Method

Such models are laborious to train, even on state-of-the-art technology; one solution is transfer learning, in which a previously-created and trained model is utilised as the basis for a new task. When it comes to transfer learning, you can take one of two common routes: Create a template method. Strategy number two: a retrained model. Selecting a task that is similar to the current work and has a lot of data is the first step in constructing a model approach. After fitting the model to the data and reaching a satisfactory level of presentation, it is used as a jumping off point for the second job of interest. With the pre-trained perfect method, you start by picking a model that has already been pre-trained. For instance, you can utilise the ImageNet dataset's pre-trained weights as a starting point for training your model on another dataset.

After obtaining the replica's pre-trained weights, the next step in implementing transfer learning is to use the earlier layers as feature extractors while discarding the fully associated layers at the top of the model that were accountable for the classification part of the dataset on which the model was trained. The model might be trained to learn the functions from the feature extractor to the output classes using the novel dataset by including classification layers. DenseNet121 pre-trained CNNs were employed in this investigation

3.4.2 DenseNet 121

DenseNets [32] are a type of convolutional neural network in which every layer is directly coupled to the one below it. The feature maps from each preceding layer are combined to form the input for the current layer. If we assume an L-layer network, then. The lth layer of the network receives as input a feature map that is a concatenation of the preceding L 1 levels. Hence there are a total of (L(L+1))/2 links in the network, one for each of the L nodes. As there is no longer a need to re-learn superfluous feature maps, this method results in a network that requires less parameter than conventional convolutional neural networks. As signal, this dense connection arrangement also mitigates the vanishing gradient problematic while training deeper constructions.

DenseNets have L layers, and each layer achieves some kind of non-linear modification. Where l is the layer index, H l (.) is a convolutional, rectified linear unit (ReLU), and batch normalisation (BN) function (Conv). As an expression of the network's dense connectedness at layer l's input, we have *eq. (15)*.

$$x_l = H_l([x_0, x_1, \ldots, x_{x-1}]) \quad (15)$$

Where x l is the layer's input and [x 0,x 1,...,x (l-1)] is the union of feature maps from layers 0 to l - 1. The authors modularized their network by separating its nodes into dense blocks (DB) and transition nodes (TB). Several dense layers (DL) made up of 11 Conv and 33 Conv layers make up a dense block. A batch normalisation layer, an 11 Conv layer, and a 22 regular pooling layer make up the transition blocks that are sandwiched among the dense blocks. DenseNet121 is a variant of network that consists of four dense chunks, each of which is made up of 6, 12, 24, and 16 dense layers

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

4. RESULTS AND DISCUSSION

Where x l is the layer's input and [x 0, x 1... x (l-1)] is the union of maps from layers 0 to l - 1. The authors modularized their network by separating its nodes into dense blocks (DB) and transition nodes (TB). Several dense layers (DL) made up of 11 Conv and 33 Conv layers make up a dense block. A batch layer, an 11 Conv layer, and a 22 average pooling layer make up the transition blocks that are sandwiched between the dense blocks. DenseNet121 is a variant of the network that consists of four dense blocks, each of which is made up of 6, 12, 24, and 16 dense layers.

4.1 Evaluation Matrices

We have compared the efficacy of different categorization models using the three widely-used metrics of ROC. The presentation of a classification model can be described with the help of a confusion matrix, which is a table used in learning. First, there are "True Positive" (T P) test results, which accurately classify whether or not a patient has heart disease; second, there are "True Negative" (T N) test results, which accurately classify whether or not a patient does not have heart disease; third, there are "False Negative" (F N) test results, which fourth, there are "False Positive" (FP) test results As far as predictions go, FN are among the worst you can make in the medical business. Accuracy is defined as for a dataset of scope n.

$$Accuracy = (TP + TN)/(TP + FP + FN + TN) \quad (16)$$

F1-Score is the harmonic mean of Recall.

$$Precision = TP / (TP + FP) \quad (17)$$

$$Recall = TP / (TP + FN) \quad (18)$$

Table 3: Analysis of Proposed Model on First Dataset

Methods	Accuracy (%)	Precision (%)	Recall	F1
SVM	83.96	83.25	0.8538	0.8430
EfficientNet	84.74	84.94	0.8302	0.8397
ResNet	83.71	81.25	0.8396	0.8258
VGGNet	90.89	89.53	0.8226	0.8574
AlexNet	90.93	89.22	0.8547	0.8730
DenseNet	93.86	92.16	0.8947	0.9078
LBOA-DenseNet	**96.92**	**94.81**	**0.9311**	**0.9395**

In the analysis of accuracy, the proposed model achieved 96.92%, the existing models such as SVM, EfficientNet and ResNet achieved nearly 83% to 84%, VGGNet, AlexNet and DenseNet achieved nearly 90% to 93%. When the models are tested with precision, recall and F1-score, the existing techniques such as SVM, EfficientNet, ResNet, VGGNet, AlexNet and DenseNet achieved nearly 82% to 92% and the proposed model achieved 94.81% of precision, 93.11% of recall and 93.95% of F1-score. The reason for better performance is that the optimal features are selected by LBOA model. The performance of projected model with existing techniques on second dataset is given in *table 4. Figure 2* and *3* presents the graphical analysis of projected model.

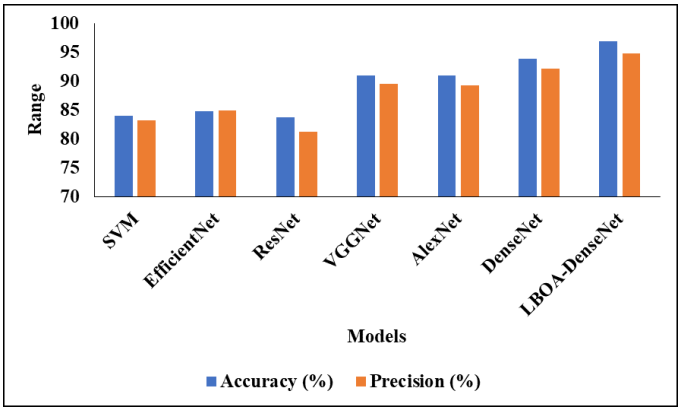

Figure 2: Comparative investigation of Proposed Model

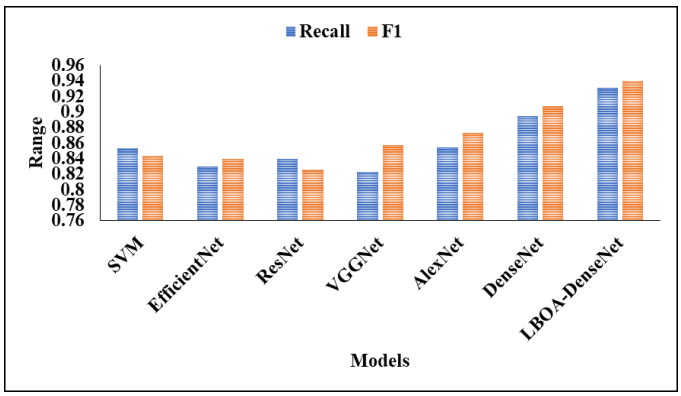

Figure 3: Analysis of Projected Model

Table 4: Analysis of Projected Model on Second Dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1
SVM	83.59	80.87	79.56	80.20
EfficientNet	82.46	82.83	80.22	81.52
ResNet	84.76	83.21	82.30	82.75
VGGNet	81.91	79.37	71.54	75.25
AlexNet	88.64	87.36	87.55	87.45
DenseNet	90.85	89.15	83.94	86.46
LBOA-DenseNet	94.92	92.87	92.98	92.92

In the analysis of F1-score, the SVM achieved 80%, EfficientNet achieved 81%, ResNet has 82%, VGGNet achieved 75%, AlexNet achieved 87%, DenseNet achieved 86% and proposed model achieved 92%. The other existing models achieved nearly 79% to 89% of accuracy, precision and recall, where the proposed model achieved 94.92% of accuracy, 92.87% of precision and 92.98% of recall. *Figure 4 and 5* provides the graphical analysis.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

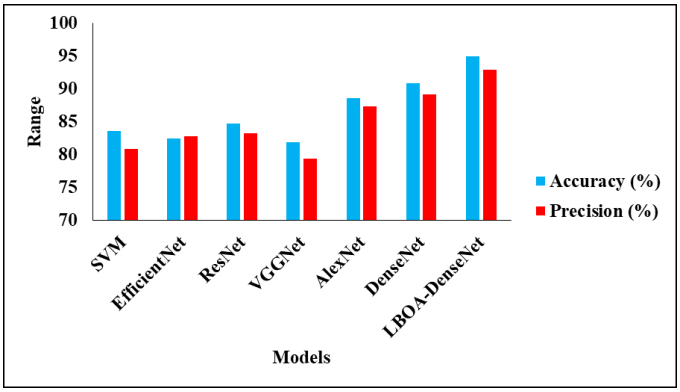

Figure 4: Analysis of Proposed model on Second dataset

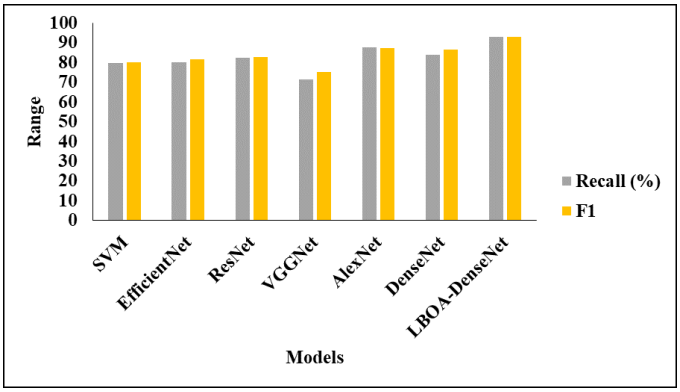

Figure 5: Performance of Various Classifier

5. CONCLUSION

The prevalence of heart disease is rising rapidly, making it one of the foremost global killers. The severity of the effects of this condition can be mitigated with prompt diagnosis and treatment. Predicting cardiac disease and choosing relevant features is the focus of this paper. The primary purpose of this investigation is to analyse how different feature selection methods affect the DenseNet model's accuracy and precision. This research was conducted on publicly available datasets for cardiovascular disease and Framingham heart disease. As a first step in our study, we pre-processed the data by transforming it, cleaning it, and balancing it. Features such as age, hypertension, and hyperglycaemia, prior were shown to indicate the most relevant risk factors for heart disease when both datasets were used for analysis. The LBOA model selects the best possible attributes. Moreover, classification experiments were conducted with both complete and reduced feature sets to examine how particular attributes affected the suggested prediction model's performance. When compared to models trained with a complete feature set, DL models performed better when their feature sets were reduced. Using a feature selection technique, the experimental findings show that we can accurately classify heart disease using fewer features and less time. The results of the research show that prediction models benefit from feature assortment since only the greatest relevant features pertaining to heart illness are considered. Our long-term goal is to develop the most effective model for cardiac disease diagnostics, and we want to do it by combining a large number of hybrid deep learning models. The research will also employ multiple feature selection methods in an effort to collect more manageable, directly applicable feature subsets for clinical research

REFERENCES

[1] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y.R. and Suraj, R.S., 2021, January. Heart disease prediction using hybrid machine learning model. In 2021 6th international conference on inventive computation technologies (ICICT) (pp. 1329-1333). IEEE.

[2] Ali, M.M., Paul, B.K., Ahmed, K., Bui, F.M., Quinn, J.M. and Moni, M.A., 2021. Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. Computers in Biology and Medicine, 136, p.104672.

[3] Jindal, H., Agrawal, S., Khera, R., Jain, R. and Nagrath, P., 2021. Heart disease prediction using machine learning algorithms. In IOP conference series: materials science and engineering (Vol. 1022, No. 1, p. 012072). IOP Publishing.

[4] Katarya, R. and Meena, S.K., 2021. Machine learning techniques for heart disease prediction: a comparative study and analysis. Health and Technology, 11, pp.87-97.

[5] Sekar, J., Aruchamy, P., Sulaima Lebbe Abdul, H., Mohammed, A.S. and Khamuruddeen, S., 2022. An efficient clinical support system for heart disease prediction using TANFIS classifier. Computational Intelligence, 38(2), pp.610-640.

[6] Rani, P., Kumar, R., Ahmed, N.M.S. and Jain, A., 2021. A decision support system for heart disease prediction based upon machine learning. Journal of Reliable Intelligent Environments, 7(3), pp.263-275.

[7] Ayon, S.I., Islam, M.M. and Hossain, M.R., 2022. Coronary artery heart disease prediction: a comparative study of computational intelligence techniques. IETE Journal of Research, 68(4), pp.2488-2507.

[8] Diwakar, M., Tripathi, A., Joshi, K., Memoria, M. and Singh, P., 2021. Latest trends on heart disease prediction using machine learning and image fusion. Materials Today: Proceedings, 37, pp.3213-3218.

[9] Nadeem, M.W., Goh, H.G., Khan, M.A., Hussain, M., Mushtaq, M.F. and Ponnusamy, V.A., 2021. Fusion-based machine learning architecture for heart disease prediction.

[10] Yazdani, A., Varathan, K.D., Chiam, Y.K., Malik, A.W. and Wan Ahmad, W.A., 2021. A novel approach for heart disease prediction using strength scores with significant predictors. BMC medical informatics and decision making, 21(1), p.194.

[11] Gao, X.Y., Amin Ali, A., Shaban Hassan, H. and Anwar, E.M., 2021. Improving the accuracy for analyzing heart diseases prediction based on the ensemble method. Complexity, 2021, pp.1-10.

[12] Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T. and Masood, M., 2021. Prediction of heart disease using deep convolutional neural networks. Arabian Journal for Science and Engineering, 46(4), pp.3409-3422.

[13] Nagarajan, S.M., Muthukumaran, V., Murugesan, R., Joseph, R.B., Meram, M. and Prathik, A., 2022. Innovative feature selection and classification model for heart disease prediction. Journal of Reliable Intelligent Environments, 8(4), pp.333-343.

[14] Magesh, G. and Swarnalatha, P., 2021. Optimal feature selection through a cluster-based DT learning (CDTL) in heart disease prediction. Evolutionary intelligence, 14, pp.583-593.

[15] Patro, S.P., Nayak, G.S. and Padhy, N., 2021. Heart disease prediction by using novel optimization algorithm: a supervised learning prospective. Informatics in Medicine Unlocked, 26, p.100696.

[16] Budholiya, K., Shrivastava, S.K. and Sharma, V., 2022. An optimized XGBoost based diagnostic system for effective prediction of heart disease. Journal of King Saud University-Computer and Information Sciences, 34(7), pp.4514-4523.

[17] Ge, B., Yang, H., Ma, P., Guo, T., Pan, J. and Wang, W., 2023. Detection of pulmonary hypertension associated with congenital heart disease based

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 253-261 | e-ISSN: 2347-470X

on time-frequency domain and deep learning features. Biomedical Signal Processing and Control, 81, p.104316.

[18] Pradhan, M. and Bhuiyan, A., 2023. Automatic Heart Disease Diagnosis Based on MRI Image Using Deep Neural Network: Adaptive Bacterial Foraging Optimization Algorithm-Based Feature Selection. In Diverse Perspectives and State-of-the-Art Approaches to the Utilization of Data-Driven Clinical Decision Support Systems (pp. 212-233). IGI Global.

[19] Hassan, D., Hussein, H.I. and Hassan, M.M., 2023. Heart disease prediction based on pre-trained deep neural networks combined with principal component analysis. Biomedical Signal Processing and Control, 79, p.104019.

[20] Nagarajan, S.M., Muthukumaran, V., Murugesan, R., Joseph, R.B., Meram, M. and Prathik, A., 2022. Innovative feature selection and classification model for heart disease prediction. Journal of Reliable Intelligent Environments, 8(4), pp.333-343.

[21] Khan, S.I., Qaisar, S.M. and Pachori, R.B., 2022. Automated classification of valvular heart diseases using FBSE-EWT and PSR based geometrical features. Biomedical Signal Processing and Control, 73, p.103445.

[22] Manimurugan, S., Almutairi, S., Aborokbah, M.M., Narmatha, C., Ganesan, S., Chilamkurti, N., Alzaheb, R.A. and Almoamari, H., 2022. Two-Stage Classification Model for the Prediction of Heart Disease Using IoMT and Artificial Intelligence. Sensors, 22(2), p.476.

[23] https://datahack.analyticsvidhya.com/contest/mckinsey-analyticsonlinehackathon/.

[24] https://inclass.kaggle.com/asaumya/healthcare-dataset-stroke-data.

[25] https://www.kaggle.com/amanajmera1/framingham-heart-study-dataset

[26] S. Sachan, F. Almaghrabi, J.-B. Yang, D.-L. Xu, Evidential reasoning for preprocessing uncertain categorical data for trustworthy decisions: An application on healthcare and finance, Expert Syst. Appl. 185 (2021) 115597.

[27] H. Wang, J. Tang, M. Wu, X. Wang, T. Zhang, Application of machine learning missing data imputation techniques in clinical decision making: taking the discharge assessment of patients with spontaneous supratentorial intracerebral hemorrhage as an example, BMC Med. Inf. Decis. Making 22 (1) (2022) 1–14.

[28] M.R. Stavseth, T. Clausen, J. Røislien, How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data, SAGE Open Med. 7 (2019) 2050312118822912.

[29] M. Buda, A. Maki, M.A. Mazurowski, A systematic study of the class imbalance problem in convolutional neural networks, Neural Netw. 106 (2018) 249–259.

[30] Arora S and Singh S 2018 Butterfly optimization algorithm: a novel approach for global optimization, Soft Comput. 715–734.

[31] Boothalingam R 2018 Optimization using lion algorithm: a biological inspiration from lion's social behavior; Evol. Intel. 11 31–52.

[32] F. Iandola, M. Moskewicz, S. Karayev, R. Girshick, T. Darrell, and K. Keutzer, ''DenseNet: Implementing efficient ConvNet descriptor pyramids,'' 2014, arXiv:1404.1869.