

# PRINCIPAL COMPONENT ANALYSIS AND HYBRID FUZZY CONVOLUTION NEURAL NETWORK TECHNIQUES BASED HEART DISEASES PREDICTION MODEL

<sup>1</sup>\*N. Vellingiri, <sup>2</sup>Dr. S. Ranjitha Kumari

<sup>1</sup>\*PhD Research Scholar, <sup>2</sup>Associate Professor

<sup>1</sup> [vellingiri.nm@gmail.com](mailto:vellingiri.nm@gmail.com), <sup>2</sup> [ranjithakumari@rvsgroup.com](mailto:ranjithakumari@rvsgroup.com)

<sup>1,2</sup>Department of computer science, RVS College of Arts and Science, Sulur, Coimbatore.

DOI: <https://doi.org/10.63001/tbs.2024.v19.i02.S2.pp163-170>

## KEYWORDS

Heart disease,  
Clinical Decision Support  
System,  
Convolutional Neural  
Network,  
Principal Component  
Analysis and  
Dove Swarm Optimization.

Received on:

10-07-2024

Accepted on:

19-10-2024

## ABSTRACT

Heart diseases cause greater mortalities globally which can be reduced by detecting them in their early stages where Clinical Decision Support Systems (CDSS) have been designed. Healthcare organizations use Hybrid Convolutional Neural Network (HCNN) algorithms to determine a patient's risk of dying or being hospitalized from heart failure. However, because of a max-pool operation, a convolutional neural network operates much more slowly. If the computer lacks a powerful GPU, CNN's consume longer times during training as they have several layers. This work presents an efficient model for predictions of cardiac illnesses. Preprocessing is initially carried out using min-max normalization. Principal Component Analysis (PCA) is used in Dimensionality Reduction. Hybrid Fuzzy Convolution Neural Network (HFCNN), based on optimized modified chicken swarm, is used to identify cardiac disorders. This work's experimental outcomes demonstrate the suggested model's higher degrees of accuracy in identifying cardiac illnesses.

## Introduction

Heart diseases include cardiovascular diseases (CVD) that remain the top most causes of death and accounts for almost 30% of global mortalities and it is predicted that the figures would cross 22 millions by 2030 where at least 121.5 million persons in the country—or over half of all adults—are afflicted by cardiovascular diseases (CVDs) according to American Heart Association's forecast. These disorders rank within the top three in Korea where 45% of deaths in 2018 were due to heart diseases [Bertsimas et al 2021; Singh and Kumar 2020].

Heart diseases are disorders which plaque the arterial wall and impede blood flows causing heart attacks or strokes. Their causes include poor diets, no or very reduced physical activities, and excessive alcohol or tobacco consumptions. Adopting good daily habits, such as reducing salts in food, having more of vegetables and fruits, regular exercises and quitting alcohol and tobacco, can progressively reduce one's risk of developing heart disease [Shanmugasundaram et al 2018; Sultana et al 2016].

Identifying high-risk individuals for heart diseases early utilizing prediction models to reduce mortality rates and improve decision-making for extra CVD preventions and treatments are advised [Barik et al 2020]. Clinicians may evaluate patients' risks for heart diseases and prescribe appropriate drugs to control risks using predictions of CDSS [Ashri et al 2021]. Studies have demonstrated the use of CDSS in multiple areas including improve clinical decisions and preventive healthcare [Rani et al 2021; Lakshmanarao et al 2021].

Clinical decisions using machine learning (ML) have recently been used in the healthcare industry, where prior studies have used Multilayer Perceptrons (MLP), Back propagation Neural

Networks (BPNN), and Chaos Firefly algorithms to assist in heart disease predictions based on individual data. Handcrafted characteristics are employed in the field of ML. Handcrafted features, on the other hand, are manual, low-level, and susceptible to limitations, such as the need for human-defined criteria that need subject-matter expertise, in contrast to automated features.

A risk model was developed for addressing the aforementioned issues with the current work. The probabilities of heart related deaths or hospitalizations for patients in a major health maintenance organization utilizing HFCNN algorithms. However, because of a maxpool operation, a convolutional neural network operates much more slowly. If the computer lacks a powerful GPU, CNN's consume longer times during training as they have several layers. This work presented an efficient model for the prediction of cardiac illnesses in order to prevent this issue. This work's principal contributions are detailed below:

- Data normalizations using min/max normalizations are executed in preprocessing steps.
- During the dimensionality reduction step, PCA is utilized to do dimensionality reduction.
- During the diagnosis stage, heart disease detection is done using HFCNN based on modified chicken swarm optimization.

### 1. Literature review

Fitriyani et al [2020] suggested CDSS based Heart disease prediction model (HDP) where XGBoost predicted CVDs using hybrid Synthetic Minority Over-sampling Technique-Edited Nearest Neighbor (SMOTE-ENN) for balanced data distributions and Density-Based Spatial Clustering of Applications with Noise

(DBSCAN) for eliminating outliers. The study used Stat log and Cleveland datasets for assessing their suggested model's efficacy when compared to prior studies and other models which included naive bayes (NB), logistic regression (LR), multilayer perceptron (MLP), support vector machine (SVM), decision tree (DT), and random forests (RF). With accuracy rates of 95.90% and 98.40% for Cleveland and Stat log datasets, the suggested model outperformed other current models. Furthermore, their Heart Disease CDSS (HDCDSS) prototype assisted physicians in determining heart disorder statuses of patients or subjects based on their present states. Consequently could help in early interventions for cardiac diseases.

Repaka et al [2019] concentrated on identifying heart diseases using prior data with their Smart Heart Disease Prediction (SHDP) model where NB to anticipated heart disease risks. Technological progress lead to notable surges in online mobile based health technologies and used common formats for gathering necessary data. Patients' risks of heart disorders were assessed using their medical profiles including blood pressures, cholesterol levels, ages, blood sugar levels, etc. The acquired attributes were input for categorization and prediction of heart diseases using NB. Two parts of the dataset were separated: training (80%) and testing (20%) where the recommended solution included the following steps: obtaining datasets, simplifying user registration and login (via an application), and categorizing data using NB, generating predictions, and securely sending data using AES. The study employed data mining techniques to anticipate heart illness to clarify and suggest several knowledge abstraction strategies. Their findings show that the present diagnostic technique accurately predicts heart disease risk factors.

Karayılan and Kılıç [2017] suggested a back-propagation algorithm-based A technique for predicting cardiac disease using artificial neural networks. Neural networks (NNs) were trained using back propagations for predicting presence/absence of heart diseases, obtaining 95% accuracy. The neural network received 13 clinical characteristics as input.

Saxena and Sharma [2016] designed a framework to predict patients' risk levels based on specified health parameters. This study's key contribution was to help non-specialist clinicians make educated judgments about r patients' heart disease risks. Original, trimmed, duplicate-free, categorized, sorted, and polish rule categories were produced by the suggested technique and framework's arrangement accuracies evaluated performances. The work's experimental results showed the framework could improve the accuracy of risk predictions in CVDs.

Nandy et al [2023] suggested an innovative healthcare system for the prediction of cardiovascular illness using Swarm based ANNs called Swarm-ANN. The suggested Swarm-ANN method starts by randomly creating a certain number of NNs and uses that number to train and evaluate the framework based on how consistently the NNs provide results. Additionally, weight alterations are applied to the NN populations during two phases of training, and

the weight modifications are made using heuristics and distributing altered global ideal weights across the neurons resulting in better precision of cardiovascular disease predictions. Using a benchmark dataset as a guide, their proposed Swarm-ANN approach predicts cardiovascular illness with 95.78% accuracy. Their simulation findings demonstrated that, for a range of performance metrics, Swarm-ANN outperformed traditional learning methods.

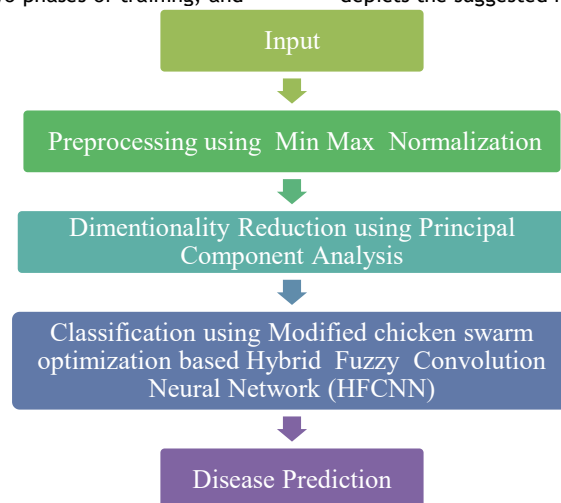
Gupta et al [2019] suggested a framework for machine intelligence (MIFH) to identify heart diseases. Features from UCI heart disease Cleveland dataset were identified for training ML prediction models where MIFH factorized mixed data, or FAMD for analysis. MIFH framework was validated using holdout validations. The experiment's findings demonstrate that, while MIFH's sensitivity and specificity were comparable to those of other previously developed baseline techniques, it outperformed them in terms of accuracy. By enhancing the effectiveness of the system and providing the most practical response out of all input prediction models while accounting for performance parameters, MIFH can help radiologists and physicians diagnose heart patients more accurately.

Long et al [2015] recommended rough sets-based attribute reductions and interval type-2 fuzzy logic systems (IT2FLS) for diagnosis of cardiac diseases. To handle the uncertainties and difficulties related to high-dimensional datasets, IT2FLS and attribute reductions using rough sets were combined where hybrid learning procedures blending fuzzy c-mean clustering, chaotic firefly parameter modifications, and genetic hybrid algorithms were used by IT2FLS. The computing cost of this learning process rises when high-dimensional datasets are employed. Rough sets-based attribute reductions utilizing chaotic firefly techniques were studied to identify best reductions and minimize computational burdens resulting in improved performances of IT2FLS. The study's experimental results showed better performances in comparison to other ML techniques including ANNs, SVM, and NB. The model could be used in decision supports for identifying cardiovascular illness.

Mehmood, et al [2021] outlined Cardio Help that used deep learning-based convolutional neural networks (CNNs) for estimating patients' risks of acquiring cardiovascular diseases. The suggested approach uses CNN to predict HF early on in order to represent temporal data. Building datasets of cardiac diseases and comparing outcomes of current methods produced good results. Their experimental outcomes with an accuracy of 97% showed better performances when compared to existing methods.

## 2. Proposed methodology

This section covers this work's suggested model in detail. The first phase normalizes data using min-max normalizations. The second involves reducing dimensionality by principal component analysis, while the third involves classifying data using an HFCNN based on modified chicken swarm optimization (CSO). Figure 1 depicts the suggested model's overall design.



**Figure 1. Proposed Flow Diagram**

**2.1. Data normalizations using min/max normalizations**

Normalizing the data is necessary to prevent erroneous findings from being produced by variances in the input crime data. This study adopts the Min-Max Normalization paradigm, which involves employing a mathematical function to transform numerical values into a new range [Yin et al 2017; Gumaei et al 2019]. The crime dataset in this proposed study is normalized using often used min-max normalizations. The dataset's values are normalized to fall between the minimum and maximum values specified, and each value is then changed using Equation (1)

$$v' = \frac{v - \min_A}{\max_A - \min_A} (\text{new\_max}_A - \text{new\_min}_A) + \text{new\_min}_A \quad (1)$$

Where,

A - Data Attributes,

Min (A), Max (A) - stands for A's min. and max. absolute values

v' - New values of inputs

v - Old values of inputs

New\_ max (A), new\_ min (A) imply max and min value ranges (boundaries of required ranges).

**2.2. Dimensionality Reduction using PCA**

On normalizing data their dimensionalities are reduced. PCA is being used in this effort to reduce dimensionality. A dimensionality reduction method called principal component

analysis creates new features by combining the original features in a linear fashion. Instances of the dataset are mapped via PCA from d-dimensional spaces to k-dimensional subspaces where  $k < d$ . The Principal Components (PC) are the collection of k newly created dimensions. Each PC aims to maximize variance while avoiding variation that has already been taken into account by all of its previous components [Du et al 2016; Tian et al 2015]. The first component, then, covers the highest variation, while each subsequent component covers a smaller degree of variance. Principal components can be computed using:

$$PC_i = a_1X_1 + a_2X_2 + \dots + a_dX_d \quad (2)$$

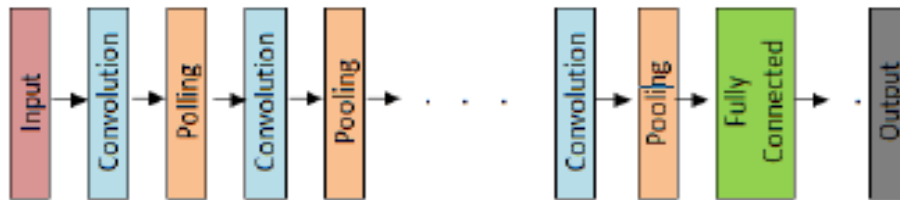
Where  $PC_i$  - Principal Component 'i';  $X_j$  - original feature 'j';  $a_j$  - numerical coefficient for  $X_j$ .

**2.3. Classification using Modified chicken swarm optimization based HFCNN**

After dimensionality reduction it sends for classification. This work using modified CSO based hybrid fuzzy CNNs.

• **CNNs**

CNN's structure differ from normal ANN structures. CNN's layers are selected to spatially suit inputs unlike a standard ANN, which flatten inputs into vectors. Tradition CNNs are made of singular or multiple subsampling and connected layers [Wu et al 2019; Gao et al 2019]. The standard design of CNN is shown in Figure 1, and the layers are explained below:



**Figure 2. Typical CNN**

• **Disadvantages of traditional CNN**

The dimension reduction process in the conventional CNN architecture involves pooling, which occasionally leads to information loss. Fuzzy CNN was employed in this work to solve this problem.

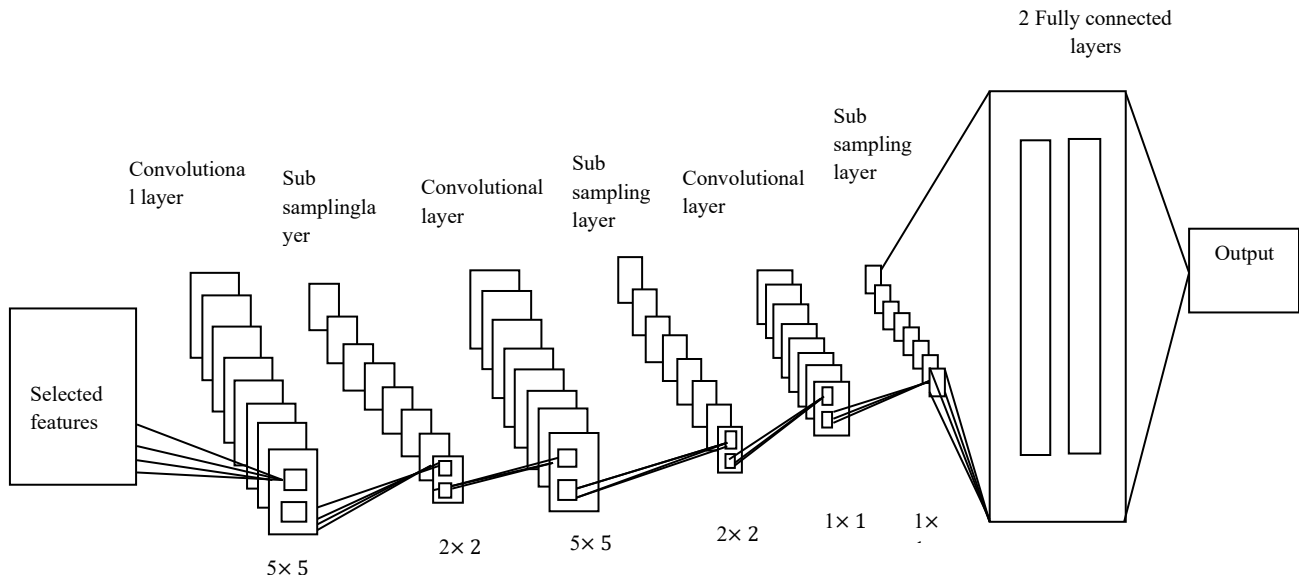
• **Fuzzy CNN**

CNN's fuzzy Three different layers i.e. fully linked, subsampling, and convolution layers form CNNs where its typical structure is depicted in. Figure 1. Subsequent sections explain these layers.

• **Convolution layer**

Kernels (filters) convolves input features in convolution layers with n output feature maps are generated from input features and results of kernels' convolutions. The outputs generate features obtained by convolving kernels and inputs in terms of feature maps with sizes  $i \times i$  which are convolution matrices and are also called filters.

A CNN can have many convolutional layers; the feature vector serves as both inputs and outputs for succeeding convolutional layers which have arrays of n filters. The feature map depths ( $n^*$ ), which are formed by convolving these filters with inputs are equal to filter counts utilized in convolution procedures where these maps can be created as unique features of input points.



### Figure.3. Convolutional Neural Network architecture

The  $l$ -th convolution layer outputs  $C_i^{(l)}$  include feature maps computed as;

$$C_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{a_i^{(l-1)}} K_{i,j}^{(l-1)} * C_j^{(l-1)} \quad (3)$$

Where,  $B_i^{(l)}$  stands for bias matrices and  $K_{i,j}^{(l-1)}$  implies kernels with  $a^*a$  sizes that connect  $(l-1)$  layers'  $j$ -th and  $i$ -th feature maps.

Outputs  $C_i^{(l)}$  result in feature maps. In (4), first convolution layers  $C_i^{(l-1)}$  are input spaces i.e.  $C_i^{(0)} = X_i$ . Activation functions are applied on these generated feature maps for nonlinear transformations of convolved layer outputs.

$$Y_i^{(l)} = Y(C_i^{(l)}) \quad (4)$$

Where,  $Y_i^{(l)}$  denote activation functions and  $C_i^{(l)}$  stand for received inputs.

#### 2.3.1. Pooling or Sub sampling Layers

These layers reduce dimensionalities of feature maps in terms of geography that were retrieved from preceding convolution

layers. Subsampling takes place between masks and feature maps. A number of subsampling strategies, such as average, sum, and maximum pooling, were suggested. Max pooling, which gives an output characteristic the greatest value of every block, is the most used type of pooling. Recall that the ability of the convolution layer to withstand picture translation and rotation is enhanced by the subsampling layer.

#### 2.3.2. Fully Connected layer

Final layers of CNN are traditional feed forward networks with multiple hidden layers where outputs use SoftMax activations:

$$Y_i^{(l)} = f(z_i^{(l)}) \quad (5)$$

$$\text{Where } z_i^{(l)} = \sum_{i=1}^{m_i^{(l-1)}} w_{i,j}^{(l)} Y_i^{(l-1)} \quad (6)$$

Where,  $w_{i,j}^{(l)}$  imply  $h$  class's tuning of weights by fully connected layers and  $f$  represents functions for non-linear transfers. It is important to note that fully connected layer's nonlinearities are built within neurons in contrast to pooling and convolutions which build it in separate layers.

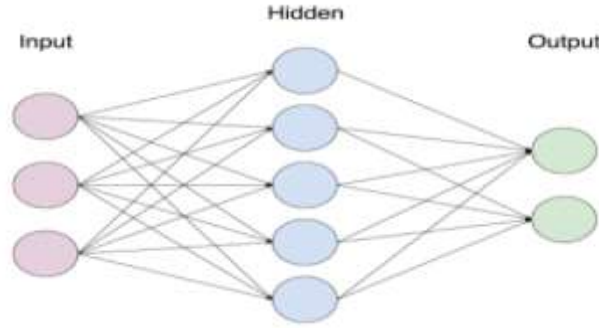


Figure4. Fully Connected Layer

Equations above compute weights by using fuzzy membership functions defined by ( $w_1 = 0.3, w_2 = 0.4, w_3 = 0.5, w_4 = 0.7$ ) and using:

$$o^2 = u_i^{(j)}(a_i^{(2)}) \quad (7)$$

Where  $u_i^{(j)}(\cdot)$  represents Gaussian membership function  $u_i^{(j)}(\cdot): \mathbb{R} \rightarrow [0, 1], i=1,2,\dots,M, j=1,2,\dots,N$ .

#### 2.3.3. CNN parameter tuning using modified chicken swarm optimization

A CNN model's parameters, such as its kernel size (KS), stride (S), and filter number (NF) (convolution layer), determine how successful the model is. This has led some academics to argue that fine-tuning these hyperparameters is essential to achieving successful outcomes. In this paper, the dove swarm optimization process is used to get the desired values.

#### 2.3.4. Chicken swarm optimization (CSO)

CSO optimizations are meta-heuristics with biological inspirations which mimics hierarchical structures of chicken swarms and individual movements of birds. Swarm hierarchical structures can be viewed as groups with a rooster in lead followed by hens and chicks. These birds follow distinct laws of movements. The social lives of chickens are significantly impacted by a hierarchical framework. Stronger chickens subjugate weaker ones. Chickens near group peripheries include both submissive hens and roosters and more dominant hens that stay near head roosters [Osamy et al 2020; Hafez et al 2015].

Local optimal characteristics are a simple trap for conventional CSO to fall into. To address this issue, this work makes use of CSO's mutation operator. In this work, flip bit mutation was used. This mutation operator flips the bits of a specified genome. (That is, the genomic bit is changed from 1 to 0, and vice versa if it is 1.

Mutation, Chicken Swarm Optimization (MCSO)

#### 2.3.5. MCSO mathematical model follows chickens' actions as detailed below:

- 1) Chicken swarms are made up of several groups. Groups with roosters in the front are followed by hens and chicks.
- 2) The fitness values of hens determine hierarchies within swarms: individual chickens are represented by the lowest fitness values, while group leaders are the roosters with the highest fitness values. Among them would be chickens.
- 3) Mother-child bonds remain unchanged in groups within swarms. Time steps (G) divide states from one another.
- 4) The four groups of the N virtual chickens that comprise the swarm are RN, HN, CN, and MN, which stand for counts of hens, chicks, roosters, and mother hens, respectively and locations in a D-dimensional spaces are represented by

$$X_{i,j} (i \in [1, \dots, N], j \in [1, \dots, D]) \quad (8)$$

**Rooster Movements:** Equations (8) and (9) demonstrate highly fit roosters hunting for food in more locations when compared to roosters with lower fitness.

$$x_{i,j}^{t+1} = x_{i,j}^t * (1 + \text{Randn}(0, \sigma^2)) \quad (9)$$

$$\sigma^2 = \begin{cases} 1, & \text{if } f_i \leq f_k, | \\ \exp\left(\frac{f_k - f_i}{|f_i| + \epsilon}\right) & \text{otherwise } k \in [1, N], k \neq i, \end{cases} \quad (10)$$

where  $x_{i,j}$  represents selected roosters with indices  $i$ ,  $\text{Randn}(0, \sigma^2)$  represents Gaussian distribution having 0 as means and standard deviations  $\sigma^2, \epsilon \in s$  signifying smallest constants in computers used to avoid divide by zero errors,  $k$  signifies rooster indices selected randomly,  $f_i$  stands for fitness values of corresponding roosters  $x_i$ .

**Hens' movements:** Hens follow roosters in pursuit of food. Moreover, they would carelessly seize the delicious food they discovered while being repressed by the other hens. In a competition for food, the more assertive chickens would have an edge over the more timid ones. Equations (10) and (11) provide a mathematical formulation for these events.

$$x_{i,j}^{t+1} = x_{i,j}^t + S1 * rRand * (x_{r_1,j}^t - x_{i,j}^t) + S2 * Rand * (x_{r_2,j}^t - x_{i,j}^t) \quad (11)$$

$$S1 = \exp((f_i - f_{r_1}) / \text{abs}(f_i) + \epsilon) \quad (12)$$

$$S2 = \exp((f_{r_2} - f_i)) \quad (13)$$

where Rand stands for uniform random values between [0, 1].  $r_1 \in [1, \dots, N]$  signifies  $i^{\text{th}}$  hens' group mates, while  $r_2 \in [1, \dots, N]$  implies roosters/hens indices in swarms while  $[1, \dots, N]$  are randomly selected indices [Ahmed et al 2017].

**Chick movement:** Chicks move around mothers in search of food and depicted as Equation (17).

$$x_{i,j}^{t+1} = x_{i,j}^t + FL * (x_{m,j}^t - x_{i,j}^t) \quad (14)$$

Where  $x_{m,j}^t$  signifies positions of  $i^{\text{th}}$  chick's mothers where  $m \in [1;N]$ , The parameter FL denotes the speed at which a chick would follow its mother; it is generated at random within the range [0, 2] to account for the variations between each chick.

Feature spaces are relatively wide and features are represented with values between 0 to 1. Finding best places in search spaces that maximize specified fitness functions, intelligent searching strategies are pertinent. The CSO's fitness function, represented by equation (15), uses the training data to minimize the feature counts chosen while maximizing classification performance throughout the validation set.

$$f_{\theta} = \omega * E + (1 - \omega) \frac{\sum_i \theta_i}{N} \quad (15)$$

where  $f_{\theta}$  represents fitness functions of vectors  $\theta$  with  $N$  denoting counts of dataset's features,  $E$  denoting classifiers' errors, and  $\omega$  represents constants regulating classification performances of selected feature counts. The 0/1 components indicate unselected / selected features.

Given dataset's feature counts correspond to used variable counts. Variables are limited to intervals [0, 1], where characteristics that are connected with their approaches 1 and can be used for classifications. The thresholds, which specify which traits in particular needs to be assessed in accordance with equation (16) are variables in individual fitness computations.

$$f_{i,j} = \begin{cases} 1 & \text{if } X_{i,j} > 0.5 \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

Where  $X_{ij}$  represents dimensions of search agents  $i$  at dimensions  $j$ . In updating firefly positions, simple truncation procedures were used to ensure variable boundaries as new values could violate limiting constraints [0, 1] for some dimensions.

1. Define starting values for RN, HN, CN, MN, and G;
2. Randomize swarm's initial hen values.
3.  $X_i$  ( $i = 1, 2, \dots, N$ );
4. Define max. iteration counts ( $T_{\text{max}}$ );
5. while  $T < T_{\text{max}}$  do for each iteration
6. if  $T \% G$  equals 0 then
7. Create hierarchical structures in swarms by ranking fitness of chickens;
8. Split swarms into groups and observe how each group's chicks and mother hens interact;
9. end
10. for each chicken  $X_i$  in the swarm do
11. if  $X_i$  is a roster then
12. Modify  $X_i$ 's location using equation 10;
13. end

14. if  $X_i$  is a hen then
15. Modify  $X_i$ 's location using equation 12;
16. end
17. if  $X_i$  is a chick then
18. Modify  $X_i$  locations with Equation (15);
19. end
20. Evaluate new solutions with Equation (16);
21. If new solutions are better than previous ones, update them;
22. end
23. end
24. Use flip bit mutations on updated solutions
25. Evaluate new solutions with Equation (16)
26. end

### 3. Results and discussion

The experimental results of the proposed model are presented in detail in this section. The suggested idea is implemented using Mat Lab. The new HFCNN model is compared with the existing CNN and Swarm-ANN models for the Cleveland database from the UCI data repository in terms of precision, recall, accuracy, and f-measure. The Cleveland database is accessible at [https://archive.ics.uci.edu/ml/machine-learning-databases/heartdisease/...](https://archive.ics.uci.edu/ml/machine-learning-databases/heartdisease/) The term "multivariate" describes this kind of dataset, which are numerical data analyses that provide or incorporate wide ranges of mathematical or statistical variables. Fourteen characteristics make up this composite: blood pressure at rest, serum cholesterol, blood sugar while fasting, maximal heart rate reached, angina brought on by exercise, old peak - ST depressions due to activities in comparison to rest - slopes of peak exercise ST segments, and counts of main vessels, and thalassemia. The Cleveland database has 920 occurrences. The performance comparison results are displayed in Table 1.

#### 3.1. Performance Metrics

##### 1) Precisions

Precisions are percentages of results which are relevant and computed using

$$\text{Precision} = \frac{\text{Truepositive}}{\text{truepositive} + \text{falsepositive}} \quad (17)$$

##### 2) Recalls

Recalls are percentages of properly classified total relevant results by the suggested algorithm and defined as

$$\text{Recall} = \frac{\text{Truepositive}}{\text{truepositive} + \text{FalseNegative}} \quad (18)$$

##### 3) Accuracies

Accuracies are fractions of predictions that models get right and defined as:

$$\text{Accuracy} = \frac{\text{Truepositive} + \text{TrueNegative}}{\text{Total}} \quad (19)$$

##### 4) F measures

F-scores are harmonic means of systems' precisions and recalls and computed as

$$2 \times [(\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})] \quad (20)$$

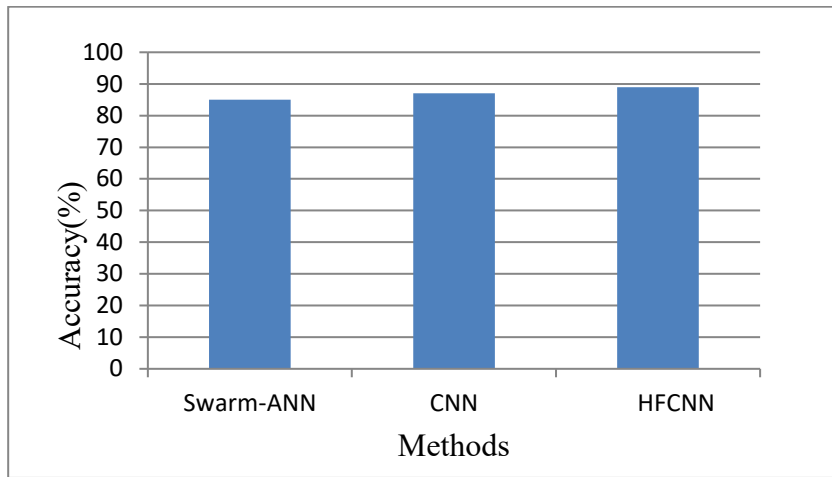
##### 5) Error rates

Error rates are measured prediction errors of models w.r.t. models and computed as

$$\text{Error rate} = 100 - \text{Accuracy} \quad (21)$$

Table.1. Performance comparison results

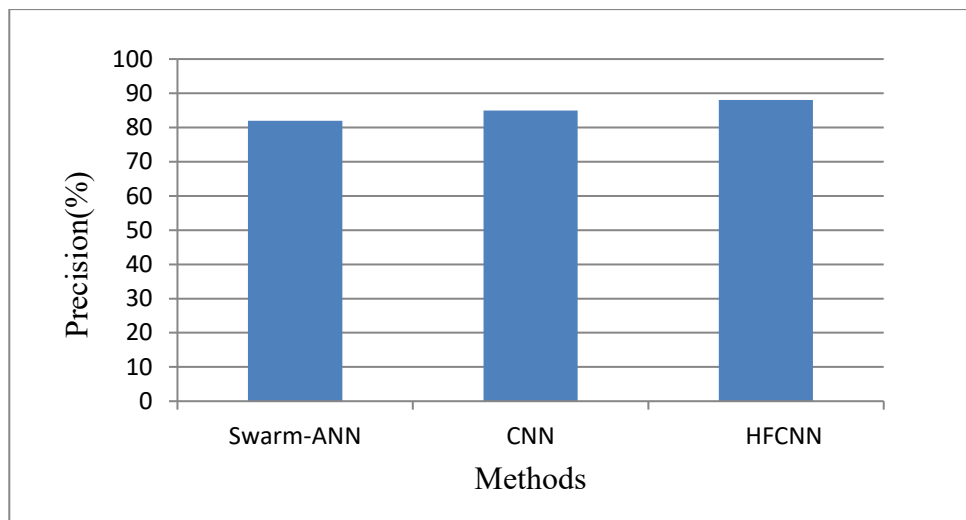
Metrics	Methods		
	Swarm-ANN	CNN	HFCNN
Accuracy (%)	85	87	89
Precision (%)	82	85	88
Recall (%)	86	87	90
F -measure (%)	87	89	92
Error rate (%)	15	13	11



**Figure.5. Accuracy results**

Figure 5 shows that in comparison to other methods, the HFCNN approach achieves excellent accuracy, meaning it can successfully identify characteristics related to heart disease in an efficient manner. As a result, the HFCNN analyzes the features in the training field and feature space and is able to effectively identify any changes in the features. In addition, the

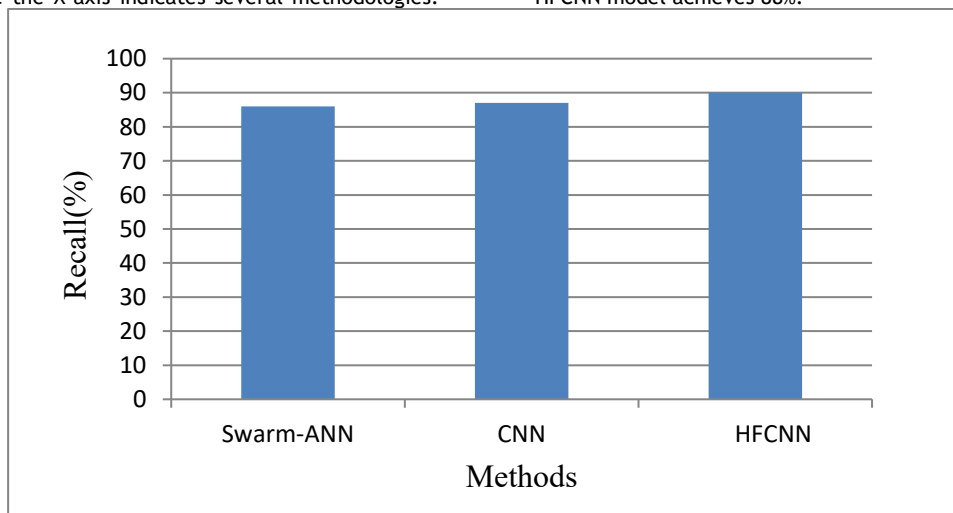
approach makes use of pertinent features that are chosen using PCA. It makes use of covariance matrices, which effectively lowers dimensionality difficulties, and it can identify heart illness with 89% recognition accuracy, compared to 85% and 87% for Swarm-ANN and CNN.



**Figure.6. Precision results**

The suggested HFCNN's performance is contrasted with that of several other algorithms, including CNN and Swarm-ANN. A graph is created to demonstrate the suggested HFCNN's good performance. The precise findings are shown on the y-axis in the above image, while the X-axis indicates several methodologies.

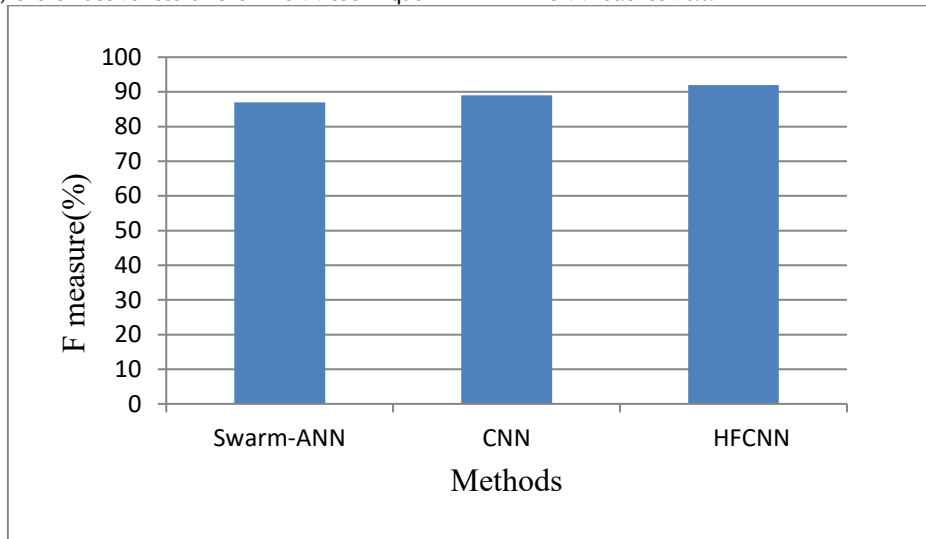
The input scale is normalized in the proposed work using the min max normalization model, which also improves the suggested HFCNN model's accuracy. According to the findings, the existing Swarm-ANN and CNN models only yield 82% and 85% of the greater precision outcomes, respectively, but the suggested HFCNN model achieves 88%.



**Figure 7. Recall results**

Figure 7. shows the outcomes of the Swarm-ANN, CNN, and HFCNN techniques' heart disease categorization recall metrics. The training set's incoming features are categorized using the HFCNN technique. When compared to traditional classifiers like Swarm-ANN and CNN, the effectiveness of the HFCNN technique

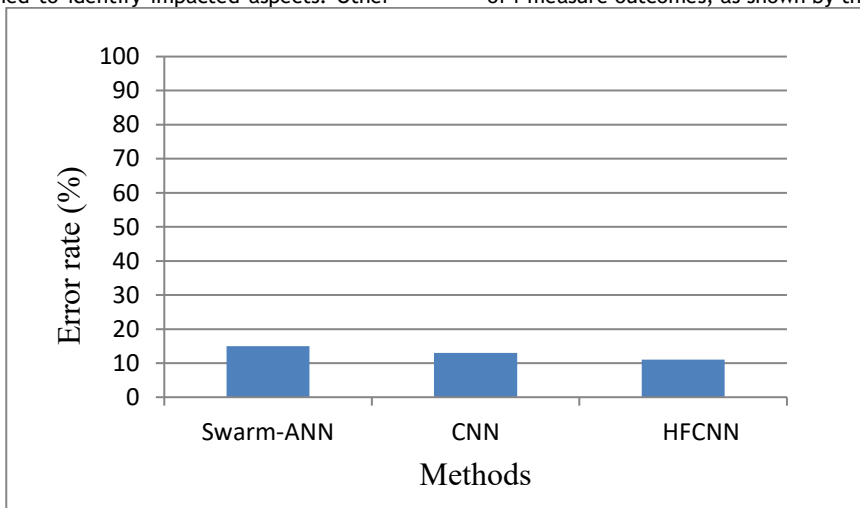
recall is greater because of the successfully picked up feature that is utilized to classify the normal and abnormal characteristics. The current Swarm-ANN and CNN models acquire recall rates of 86% and 87%, respectively, whereas the proposed HFCNN reaches 90%.



**Figure 8. F-measure results**

Figure. 8 displays the f measure as a function of the suggested HFCNN algorithm and the current CNN and Swarm-ANN techniques. The suggested HFCNN showed a dependable result that may be applied as real smart technology in medical data mining after it was trained to identify impacted aspects. Other

research investigated whether the capacity to discriminate amongst ML-based models is unique. HFCNN scores 92%, Swarm-ANN scores 87%, and CNN scores 89% on the chosen attributes. The suggested HFCNN model outperforms other models in terms of f measure outcomes, as shown by the above image.



**Figure 9. Error rate results**

In figure. 9, HFCNN guarantees an extremely low error rate of 11%, which is far lower than competing methods like CNN and Swarm-ANN. Models like Swarm-ANN and CNN have error rates of 15% and 13%, respectively. A high classification rate for heart disease prediction is achieved by the effective update of network weights, bias, and learning processes, which reduce

error rates when measured using precision and recall metrics. Table 1 displays the obtained value. During the training phase, incoming inputs are analyzed using several perceptions that are then utilized to effectively classify the testing data. A good training procedure lowers the error rate during the classification phase.

**4. Conclusion and future work**

max normalization is ineffective at handling outliers, hence alternative normalization models will be required in the future.

In the modern world, heart disorders are a major cause of early disability and death. This suggests that their prognoses are serious problems for the healthcare systems. This study advances serious knowledge and advances the creation of an intelligent system based on the DL technique for the prediction of heart disease. The foundation for data normalization in this work is min-max normalization. In this study, PCA is used to reduce dimensionality and to identify cardiac issues, this paper uses a hybrid fuzzy convolution neural network (HFCNN) based on modified chicken swarm optimization. The experimental findings reveal that the suggested model has a high accuracy of 89%. Min

**References**

- Bertsimas, D., Mingardi, L. and Stellato, B., 2021. Machine learning for real-time heart disease prediction. IEEE Journal of Biomedical and Health Informatics, 25(9), pp.3627-3637.
- Singh, A. and Kumar, R., 2020, February. Heart disease prediction using machine learning algorithms. In 2020 international conference on electrical and electronics engineering (ICE3) (pp. 452-457). IEEE.

- Shanmugasundaram, G., Selvam, V.M., Saravanan, R. and Balaji, S., 2018, July. An investigation of heart disease prediction techniques. In 2018 IEEE International Conference on System, Computation, Automation and Networking (ICSCA) (pp. 1-6). IEEE.
- Sultana, M., Haider, A. and Uddin, M.S., 2016, September. Analysis of data mining techniques for heart disease prediction. In 2016 3rd international conference on electrical engineering and information communication technology (ICEEICT) (pp. 1-5). IEEE.
- Barik, S., Mohanty, S., Rout, D., Mohanty, S., Patra, A.K. and Mishra, A.K., 2020. Heart disease prediction using machine learning techniques. In *Advances in Electrical Control and Signal Systems: Select Proceedings of AECSS 2019* (pp. 879-888). Springer Singapore.
- Ashri, S.E., El-Gayar, M.M. and El-Daydamony, E.M., 2021. HDPF: heart disease prediction framework based on hybrid classifiers and genetic algorithm. *IEEE Access*, 9, pp.146797-146809.
- 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.
- Lakshmanarao, A., Srisaila, A. and Kiran, T.S.R., 2021, February. Heart disease prediction using feature selection and ensemble learning techniques. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 994-998). IEEE.
- Fitriyani, N.L., Syafrudin, M., Alfian, G. and Rhee, J., 2020. HDPM: an effective heart disease prediction model for a clinical decision support system. *IEEE Access*, 8, pp.133034-133050.
- Repaka, A.N., Ravikanti, S.D. and Franklin, R.G., 2019, April. Design and implementing heart disease prediction using naive Bayesian. In 2019 3rd International conference on trends in electronics and informatics (ICOEI) (pp. 292-297). IEEE.
- Karayılan, T. and Kılıç, Ö., 2017, October. Prediction of heart disease using neural network. In 2017 international conference on computer science and engineering (UBMK) (pp. 719-723). IEEE.
- Saxena, K. and Sharma, R., 2016. Efficient heart disease prediction system. *Procedia Computer Science*, 85, pp.962-969.
- Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V.G., Li, X. and Zakarya, M., 2023. An intelligent heart disease prediction system based on swarm-artificial neural network. *Neural Computing and Applications*, 35(20), pp.14723-14737.
- 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.
- Long, N.C., Meesad, P. and Unger, H., 2015. A highly accurate firefly based algorithm for heart disease prediction. *Expert Systems with Applications*, 42(21), pp.8221-8231.
- 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.
- Yin, Z., Wan, B., Yuan, F., Xia, X. and Shi, J., 2017. A deep normalization and convolutional neural network for image smoke detection. *IEEE Access*, 5, pp.18429-18438.
- Gumaei, A., Hassan, M.M., Hassan, M.R., Alelaiwi, A. and Fortino, G., 2019. A hybrid feature extraction method with regularized extreme learning machine for brain tumor classification. *IEEE Access*, 7, pp.36266-36273.
- Du, M., Ding, S. and Jia, H., 2016. Study on density peaks clustering based on k-nearest neighbors and principal component analysis. *Knowledge-Based Systems*, 99, pp.135-145.
- Tian, J., Morillo, C., Azarian, M.H. and Pecht, M., 2015. Motor bearing fault detection using spectral kurtosis-based feature extraction coupled with K-nearest neighbor distance analysis. *IEEE Transactions on Industrial Electronics*, 63(3), pp.1793-1803.
- Wu, H., Chen, J., Liu, X., Xiao, Y., Wang, M., Zheng, Y. and Rao, Y., 2019. One-dimensional CNN-based intelligent recognition of vibrations in pipeline monitoring with DAS. *Journal of Lightwave Technology*, 37(17), pp.4359-4366.
- Gao, R., Xia, H., Li, J., Liu, D., Chen, S. and Chun, G., 2019, November. DRCLR: Deep reinforcement learning framework incorporating CNN and GAN-based for interactive recommendation. In 2019 IEEE international conference on data mining (ICDM) (pp. 1048-1053). IEEE.
- Osamy, W., El-Sawy, A.A. and Salim, A., 2020. CSOCA: Chicken swarm optimization based clustering algorithm for wireless sensor networks. *IEEE Access*, 8, pp.60676-60688.
- Hafez, A.I., Zawbaa, H.M., Emary, E., Mahmoud, H.A. and Hassanien, A.E., 2015, November. An innovative approach for feature selection based on chicken swarm optimization. In 2015 7th international conference of soft computing and pattern recognition (SoCPar) (pp. 19-24). IEEE.
- Ahmed, K., Hassanien, A.E., Ezzat, E. and Tsai, P.W., 2017. An adaptive approach for community detection based on chicken swarm optimization algorithm. In *Genetic and Evolutionary Computing: Proceedings of the Tenth International Conference on Genetic and Evolutionary Computing*, November 7-9, 2016 Fuzhou City, Fujian Province, China 10 (pp. 281-288). Springer International Publishing.