Received: 25 July 2023 | Revised: 23 September 2023 | Accepted: 23 October 2023 | Published online: 27 October 2023

RESEARCH ARTICLE

Medinformatics
2023, Vol. 00(00) 1–12
DOI: 10.47852/bonviewMEDIN32021377

BON VIEW PUBLISHING

# Enhancing Heart Disease Detection Using Political Deer Hunting Optimization-Based Deep Q-Network with High Accuracy and Sensitivity

**Chetan N. Aher**[1,*]

[1]*AISSMS Institute of Information Technology, India*

**Abstract:** Heart disease is a clinical syndrome generally caused due to the impairment of structural and functional tissues. According to the World Health Organization, heart disease has been recognized as the major reason for death occurring worldwide and the survey reported that 23% of death in the USA is mainly because of heart-related diseases. Therefore, quick and precise detection of heart-related diseases is significant to enhance the health issues of heart disease-affected people and also reduces the mortality rate of the patients by protecting them from dangerous death by offering better medical services. Nowadays, deep learning techniques play a significant part in the domain of medical science and are also utilized to diagnose various diseases. Due to poor feature representation and some imbalance problems, predicting heart disease is very difficult using simple models. To encounter such problems, an effective strategy is employed in this research for detecting heart disease using the proposed political deer hunting optimization (PDHO) algorithm-based Deep Q-network. However, the proposed PDHO algorithm is derived by the integration of the political optimizer and deer hunting optimization algorithms. Moreover, the proposed PDHO-based Deep Q-network for heart disease detection achieved a maximum accuracy of 93.4%, maximum sensitivity of 96.2%, and maximum specificity of 89.2%. This study leverages a diverse and extensive dataset from the gene expression omnibus dataset [GSE98583] to advance our understanding of coronary artery disease.

**Keywords:** heart disease, gene expression data, Deep Q-network, deer hunting optimization (DHO) algorithm, political optimizer (PO)

## 1. Introduction

Genes represent a uniform way of various expression extents in every disorder, and more complicated diseases, such as diabetes, cancer, and cardiovascular, cause extreme effects on human well-being. Such defects are the outcome of complex interactions of different genes, detecting the disease patient genes is most significant to understand the type of disease and to find better treatment ways (Aher & Jena, 2023a, 2023b). Several techniques have been developed to detect the disease patient genes depending on various categories of biological information, like function, sequence-related features, and network (Han et al., 2018; Nikdelfaz & Jalili, 2018). However, the existing methods of cardiac biomarkers consider the domain knowledge of physiological and pathological needs, which are the key elements, while the microarray fields only assume the expression of massive of genes simultaneously and concentrate on stimulating gene expression profiling over different pathways. Gene expressions normally allow us to detect and identify informative biomarkers that can reflect cardiovascular disease (CVD). Many researchers have provided significant results from this process (Khan, 2020; Neelima & Babu, 2020). According to statistics reported by World Health Organization, one-third of the population is died due to heart diseases, and it is considered as the major source of overall mortality rate in developing countries. It is proven that one-third of adults in the USA have suffered from heart disease according to the report released by American Heart Association. Computational biology is frequently utilized in the mechanism of converting biological command into clinical application and also in understanding biological phenomena from clinical information (Chen et al., 2011; Semmlow & Rahalkar, 2007).

In recent years, coronary angiography has emerged as a pivotal technology for diagnosing coronary artery disease (CAD). However, its invasive and costly nature restricts its use to specific clinical scenarios (Li et al., 2020). CAD, also known as ischemic heart disease, manifests in severe symptoms like chest pain, sudden heart attacks, and even cardiac arrest, making it a leading global cause of mortality (Li et al., 2020; Mehmood et al., 2021). Astonishingly, CAD's impact on mortality surpasses that of many tumors and other disorders, contributing to 40% of all heart-related deaths globally. The prevalence of CAD affects approximately 7.35 per million individuals (Ali et al., 2020; Candemir et al., 2020).

CAD arises from plaque accumulation in arteries, causing narrowing and hardening, leading to angina and myocardial infarctions. Severe stenosis can trigger cardiac arrest (hypoxia)

*Corresponding author: Chetan N. Aher, AISSMS Institute of Information Technology, India. Email: chetan.aher07@gmail.com

(Nikdelfaz & Jalili, 2018; Padmanabhan & Semmlow, 1994). Plaques come in three types: calcified, mixed, and non-calcified, with varying symptoms depending on type and extent (Guo et al., 2019; Xiao et al., 2020).

Recent medical advancements leverage deep learning and machine learning, enhancing therapeutic approaches. Bayesian classifiers streamline clinical workflows, while deep learning excels in clinical image analysis. Deep learning effectively detects physiological anomalies in real-time heart rate data. Genetic ensembles predict arrhythmia, and convolutional neural networks (CNNs) improve medical image analysis.

Traditional methods like multilayer perceptron, logistic regression (LR), and support vector machine (SVM) also detect heart disorders. Deep learning's prowess in handling class imbalance and feature representation issues makes it invaluable in medical decision-making. Deep CNNs extract high-quality features from neighboring data, and customized loss functions enhance model performance (Rairikar et al., 2017; Wang et al., 2020).

The main objective of this article is to introduce and develop an efficient mechanism for predicting heart disease using the newly proposed political deer hunting optimization (PDHO)-based Deep Q-network. The approach comprises three crucial steps: pre-processing, feature fusion, and disease prediction. Initially, the gene expression data are subjected to pre-processing, where a mechanism based on Box–Cox transformation is employed. After the pre-processing stage, feature fusion is carried out to combine the features using a correlation-based Deep Q-network. Subsequently, the prediction of heart disease is accomplished through the trained Deep Q-network, utilizing the developed PDHO algorithm. Notably, PDHO is derived by incorporating partial optimization (PO) and differential harmony optimization (DHO) techniques.

The primary contribution of the research revolves around the introduction of a novel heart disease detection mechanism, known as the PDHO-based Deep Q-network. This innovative approach focuses on achieving accurate predictions by employing a deep learning classifier trained with the PDHO algorithm. Through feature fusion facilitated by a correlation-based Deep Q-network, the research successfully attains a high level of accuracy in heart disease prediction. The article is structured to provide a comprehensive understanding of this contribution. It begins with an introduction that the motivation behind establishing an effective strategy for heart disease detection and conducts a thorough literature review to assess the advantages and limitations of traditional methods. Section 2 elaborates on the developed PDHO-based Deep Q-network, providing insights into its design and functioning methods and materials offer a detailed explanation of the training algorithm crucial for the network's performance. In Section 3, the research presents and discusses the results obtained from this innovative approach, emphasizing its efficacy. Finally, Section 4 wraps up the article with conclusions drawn from the study's findings and outlines potential future directions for further research in the field of heart disease detection.

By following this organized structure, the research aims to comprehensively present the novel approach for heart disease detection using the PDHO-based Deep Q-network, showcase the effectiveness of the training algorithm, and provide valuable insights from the results and discussion. Conventional approaches related to deep learning-based heart disease detection highlight their respective advantages and limitations. The shortcomings of these traditional methods will serve as motivation for the researchers to design and develop an effective deep learning-based heart disease detection method utilizing gene expression data.

A. Conventional Approaches for Heart Disease Detection:
   (a) Traditional Machine Learning: Conventional machine learning algorithms, such as SVMs, random forest, and LR, have been applied for heart disease detection. These methods use hand-engineered features and require manual feature extraction, which can be time-consuming and may not fully capture the complexity of heart disease patterns.
   (b) Rule-based Systems: Rule-based systems rely on predefined rules to identify heart disease patterns in data. These methods may lack flexibility and might not adapt well to variations in the data, potentially leading to reduced accuracy.
   (c) Expert Systems: Expert systems employ knowledge bases created by domain experts to guide decision-making in heart disease detection. However, these systems heavily rely on the accuracy and completeness of the knowledge base and may struggle to handle uncertain or ambiguous cases.

B. Advantages of Deep Learning-Based Approaches:
   (a) Automatic Feature Learning: Deep learning models can automatically learn relevant features from raw data, eliminating the need for manual feature engineering. This allows the models to capture complex patterns and relationships in the data more effectively.
   (b) Nonlinear Representations: Deep learning techniques can model nonlinear relationships present in heart disease data, which is particularly beneficial when dealing with intricate and subtle patterns that traditional linear models might overlook.
   (c) Scalability: Deep learning models are scalable and can handle large datasets efficiently, making them suitable for analyzing extensive gene expression data and capturing subtle genetic variations associated with heart disease.

C. Limitations of Conventional Approaches:
   (a) Data Representation: Conventional methods may struggle to efficiently represent high-dimensional gene expression data, potentially leading to the loss of important information during feature extraction.
   (b) Overfitting: In some cases, traditional machine learning models can be prone to overfitting, especially with limited data, resulting in reduced generalization performance.
   (c) Interpretability: Rule-based and expert systems might lack interpretability, making it challenging to understand how certain decisions are reached, limiting the trust and acceptance of these approaches in medical applications.

Various existing works of deep learning-based heart detection techniques are explained as follows: Li et al. (2020) developed a features fusion paradigm that merged both the features of hand-crafted and deep learning of phonocardiogram (PCG) signals to identify CAD. Here, the features of deep learning were excerpted from mel-frequency cepstral coefficients images using CNN. Moreover, this method utilized multiple features for better prediction and the results have proven that the performance was higher than that of hand-crafted as well as deep learning features. The only drawback exists in this method was very high time consumption during processing. In addition, it failed to combine the features reflecting dynamic information and failed to use PCG signals as it enhances the prediction accuracy of CAD. Xiao et al. (2020) presented a U-net convolutional neural classifier for the segmentation of coronary artery to predict the risk of disease and the method was very suitable for multiple datasets with two backgrounds, such as in the presence of centerline and also in the absence of centerline. The major benefits of this technique were

utilization of any size of the image and improved efficiency during segmentation. Though the results obtained were fuzzy and smooth and the results were not susceptible to information in the image, it was not adaptable to the elaborative form of medical images. Sharma and Parmar (2020) introduced a model based on deep neural network (DNN), which was utilized to enhance the overall quality of heart disease classification. Here, the heart disease UCI dataset was utilized to demonstrate Talos hyperparameter optimization, which was proven to be more effective in enhancing the accuracy of prediction. However, the method was not suitable to predict other type of diseases. Ashraf et al. (2019) developed an automatic system to predict heart attack using DNN and also effectively removed all the anomalies from the model, such as lack of accuracy. The method was tested on multiple datasets to identify true potential and provided certainty in the prediction of disease. The minimum accuracy provided by this automatic system is still a major drawback. However, the time consumption of this method was very low that provided results in a very short time.

Mehmood et al. (2021) modeled a method called CardioHelp, which was utilized to detect the probability of the existence of CVD by integrating a deep learning algorithm named CNN. The database utilized in this model consisted of both cardiac test parameters and general human habits. The classification done by this method provided high accuracy, but it failed to predict the happening of various major disorders, such as cancer and other brain-related disorders. Candemir et al. (2020) developed a deep 3-dimensional CNN (3D-CNN) that was employed to predict the pathological variances in coronary vessels. The developed model learned discriminative features between vessels with and without atherosclerosis, such that features were utilized to provide visual clues associated with atherosclerosis. The processing speed of this method was high, but the method was not effective at all the time, especially in the presence of noise. Wang et al. (2020) designed a deep learning structure called feature re-arrangement-based deep learning system (FRDLS) for detecting heart failure depending on real-world data. Two approaches, namely FReaConv and Focal loss, were employed in this architecture that assisted to provoke the prediction accuracy of heart failure mortality. The developed system was quick and accurate in detecting heart failure mortality and also effectively handled the imbalance situations in medical data. However, it failed to implement multi-label, multi-class, and other prediction tasks. Abdeltawaba et al. (2020) introduced a fully CNN (FCN) for automatic quantification of left ventricle function. The developed system provided low error for computed parameters to diagnose cardiac diseases very accurately. Moreover, the developed system had sufficient potential to automate the analysis of the cardiac functional analysis process, but the computational overhead of this method was high.

## 2. Major Challenges Observed in Various Methods

The challenges faced by conventional deep learning-based heart disease detection techniques are described as follows:

A. CNN Method (Li et al., 2020):

Very promising results were achieved for non-invasive CAD detection, but the method lacked the ability to combine features reflecting dynamic contents, potentially limiting its prediction accuracy for CAD. The method did not utilize multi-channel PCG signals, which could enhance CAD prediction accuracy but were not incorporated in the approach.

B. DNN (Ashraf et al., 2019):

The DNN method did not consistently achieve better results, presenting a challenge in ensuring consistent accuracy. To improve accuracy, the system could potentially benefit from the integration of another deep learning algorithm.

C. Deep CNN (Mehmood et al., 2021):

The overall accuracy attained by the deep CNN method was sufficient for heart disease detection, but it was not adaptable for predicting other diseases such as cancer or brain-related disorders.

D. Deep 3D-CNN (Wang et al., 2020):

The deep 3D-CNN method achieved quick results in predicting heart failure mortality and handled imbalance issues in medical data effectively. However, the method faced challenges in addressing multi-class, multi-label, and various prediction tasks.

E. FCN (Abdeltawab et al., 2020):

The FCN deep learning framework facilitated the automatic analysis process of cardiac function, but it did not integrate a computer-aided diagnostic model for heart disease detection and classification.

These challenges underscore the need for further research and refinement of deep learning-based techniques to overcome limitations and improve the accuracy, versatility, and applicability of heart disease detection models. Addressing these challenges will contribute to the development of more robust and comprehensive solutions for heart disease diagnosis and treatment.

## 3. Method and Materials

The primary aim of this research is to introduce and establish the PDHO-based Deep Q-network as a novel approach for heart disease detection. The process begins by taking gene expression data as input, which undergoes pre-processing through Box–Cox transformation. After pre-processing, feature fusion is conducted using a correlation-based Deep Q-network. The disease prediction is performed using the Deep Q-network, which is trained using the developed PDHO algorithm. The PDHO algorithm is a combination of the political optimizer (PO) (Askari et al., 2020) and deer hunting optimization (DHO) (Brammya et al., 2019) algorithms.

The Proposed PDHO Deep Q-network for heart disease detection presented in Figure 1 introduces an approach to the early detection of heart diseases. The PDHO (PO + DHO) algorithm is a unique combination for the study, whereas deep learning model is in the form of a Deep Q-network and gene expression data.

However, before analysis can begin, a critical step is undertaken pre-processing of the dataset. Through the application of the Box–Cox transformation, the gene expression data are meticulously prepared for subsequent analysis. This statistical technique is instrumental in stabilizing variance and ensuring that the data conform to a normal distribution, a crucial requirement for many machine learning models.

The most intriguing aspect of this research is the concept of feature fusion. Using a correlation-based Deep Q-network, the system effectively amalgamates various features within the gene expression data, aiming to capture complex interrelationships and correlations that might otherwise remain hidden. This step underscores the profound potential of deep learning in uncovering hidden patterns and connections in biological data.

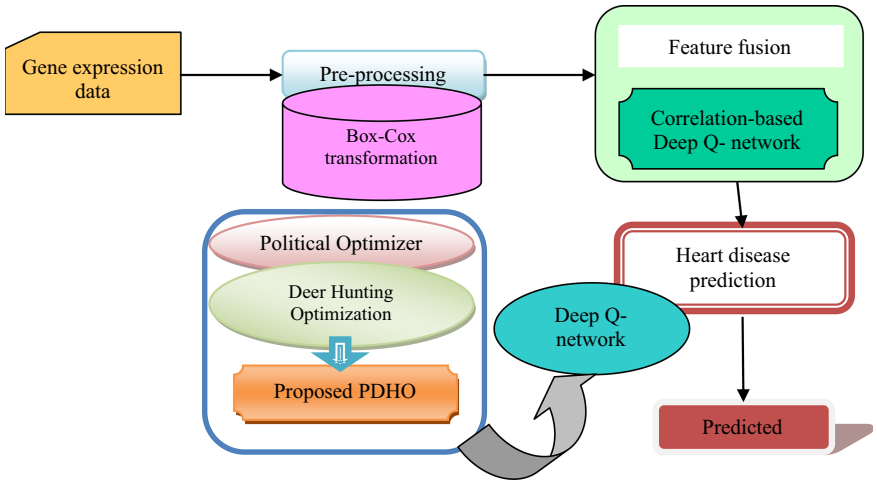


**Figure 1. Block diagram of proposed PDHO-based Deep Q-network for heart disease detection**

The major step of the system lies in the Deep Q-network, a reinforcement learning model specifically designed for decision-making and prediction tasks. In this context, it serves as the engine for heart disease detection, making informed decisions based on the knowledge it gleans from the pre-processed gene expression data. What sets this research apart is the utilization of the PDHO algorithm, a fusion of two powerful optimization techniques –PO and DHO. These optimization algorithms are harnessed to fine-tune the Deep Q-network, ensuring it operates at peak efficiency and delivers accurate and timely heart disease predictions.

## 3.1. Acquisition of gene expression data

Gene expression data, often represented as a matrix of gene activity levels across various samples, is a powerful input for predicting cardiac diseases. Its unique ability to unveil gene regulation, metabolic networks, and underlying disease mechanisms makes it an invaluable resource. These data play a crucial role in facilitating early and accurate disease diagnosis by identifying biomarkers and signatures that signify the presence and progression of cardiac conditions. By harnessing these insights, healthcare professionals can improve diagnostic accuracy, implement timely interventions, and ultimately enhance patient outcomes in the realm of cardiovascular health.

In this study, we consider a training set of gene expression data comprising a specific number of samples. The training set is expressed as follows:

$$D = \{G_1, G_2, \ldots G_i, \ldots G_n\} \tag{1}$$

where $G_i$ denotes the input data with the dimension of $[18 \times 22277]$, and the total count of samples in the training set is denoted as $G_n$.

## 3.2. Pre-processing using Box–Cox transformation

The input data $G_i$ are fed through the pre-processing step, which is a significant step before considering any deep learning techniques. Typically, it is not feasible to employ real-time data directly in the detection process because there may be a chance of generating noisy, inconsistent, and incomplete data. Hence, pre-processing phase is significant to represent data efficiently. It provides complete, clean, and accurate data, reducing computational complexity and enhancing the data's efficiency and accuracy. Generally, data pre-processing includes normalization, missing-data filtering, feature weighting, and feature selection. The obtained pre-processed result is represented $P_i$ with a dimension of $[288 \times 22277]$. Here, pre-processing is done exploiting Box–Cox transformation, which is elaborated in the below section.

### *3.2.1. Box–Cox transformation*

The transformation developed by Box and Cox (Maciejewski et al., 2012) is a specific family of power transformations with many benefits, like the transformation of data to an approximate normal distribution and stabilization of variance. Let us consider a vector with $m$ observations $y = \{y_1, \ldots, y_m\}$ and the data are converted utilizing the method of Box–Cox transformation and it is expressed by,

$$y^{(\eta)} = \begin{cases} \dfrac{y^{\eta} - 1}{\eta} & ; \eta \neq 0 \\ \ln(y) & ; \eta = 0 \end{cases} \tag{2}$$

The vector of observed or recorded data is expressed as $y$ and $\eta$ represents the power. In the case of negative data, it can be transformed to this form by including a constant.

## 3.3. Feature fusion using correlation-based Deep Q-network

Fusion is a mechanism of combining various sources of data to produce accurate and valuable data, which are utilized further for classification and detection purposes. Typically, there are three levels of fusion available, such as data level, feature level, and decision level fusion. The first type of fusion that is data level merges various data from different sources that match each other, and data fusion is broadly divided into two levels, such as feature-level fusion and decision-level fusion. At the feature level, features are reconstructed from various databases and after that combined to generate the best feature set for detection while, at the decision level, the decisions of different mechanisms are represented to improve the accuracy of the model. Generally, data-level fusion consists of an abundant quantity of unnecessary information and hence it is not feasible, whereas feature-level fusion consists of enough amount of information to predict the risk of disease.

*3.3.1. Sorting features based on correlation*

The pre-processed result is applied to the feature fusion process, where features are sorted out initially based on correlation coefficient and then grouping of features is done. Correlation coefficients (Benesty et al., 2009) are utilized to learn the strength and direction of the linear relationships between pair of variables and also used to state the degree to which two variables are associated. Moreover, features with high correlation are highly dependent and thus possess the same result as the dependent variable. Therefore, if two features possess a large correlation, then any one of the features is dropped. Thus, sorting is accomplished this way using correlation.

*3.3.2. Grouping features*

Once the features are sorted out, grouping or fusing the features is performed based on the sort-out features and is represented as,

$$K^N = \sum_{i=1}^{b} \frac{\beta}{p}, i = i + \frac{N}{g} \tag{3}$$

where $N$ denotes the total count of features and selected features are indicated as $b$. Here, $g$ specifies the ratio of the total count of features to the features to be selected, which is given by,

$$g = \frac{N}{b} \tag{4}$$

***(i) Training based on parameter $\beta$***

To determine the parameter $\beta$, Deep Q-network is utilized and ground truth is exploited to train the parameter $\beta$ and it is expressed as,

$$\beta = Correlation(DD_i, \alpha_i) \tag{5}$$

where $DD_i$ denotes the data, and the argument of the data belonging to a class is specified as $\alpha_i$. The result obtained after the completion of feature fusion is denoted as $K^N$.

***(ii) Architecture of Deep Q-network***

Deep Q-network (Sasaki et al., 2017) utilizes the Q-learning, which is recognized as the most commonly utilized reinforcement learning and exploits CNN to average the action-value function known as the Q-function. The instability condition of reinforcement learning is due to the utilization of nonlinear function approximators like neural networks to illustrate the Q-function. A Deep Q-network is a multi-layered neural network that is designed for a state that Sresults in a vector of action function $Q(S, e; \lambda)$. Here, the parameters of the neural network are denoted as $\lambda$. Figure 2 illustrates the architecture of the Deep Q-network. The loss function is represented as,

$$L_h(\lambda_h) = B_{S,e,S'}[(z_h - Q(S, e, \lambda_h))^2] \tag{6}$$

where

$$z_h = (a_h + \sigma \, max_{e'} Q(S', e', \lambda^-)) \tag{7}$$

Here, $L_h(\lambda_h)$ indicates the expected error is the parameter is $\lambda_h$. The functions of a separated target network and parameters of an online network are represented as $\lambda^-$ and $\lambda_h$, respectively. The purpose of the target network is to enhance the stability of the learning updates. However, the gradient descent is expressed as,

$$\nabla \lambda_h L_h(\lambda_h) = B_{S,e,S'}[(z_h - Q(S, e, \lambda_h)) \nabla \lambda_h Q(S, e)] \tag{8}$$

To eliminate the correlated updates, the Deep Q-network presents an experience replay with invariant maximum capacity. Transitions from earlier steps are sampled from replay memory several times to update the network; thus, the divergence problems resulting from the correlated updates are eliminated. The result obtained from Deep Q-network is specified as $T_i$.

## 3.4. Heart disease detection using Deep Q-network

After the feature fusion process, the obtained result is subjected to heart disease detection using the Deep Q-network. The Deep Q-network is responsible for identifying the presence of heart disease and categorizing the condition as either normal or abnormal. To train the Deep Q-network, we utilize the proposed PDHO algorithm, which is a novel combination of the PO (Askari et al., 2020) and DHO (Brammya et al., 2019) algorithms. This innovative algorithm enhances the training process of the Deep Q-network, leading to improved accuracy and efficiency in heart disease detection.

*3.4.1. Training procedure of proposed PDHO algorithm*

The DHO algorithm is a powerful meta-heuristic optimization technique that draws inspiration from the hunting behavior of hunters when pursuing deer. In the quest to hunt deer effectively, hunters employ various strategies, and these strategies are translated into algorithmic parameters. Some of the key parameters considered in

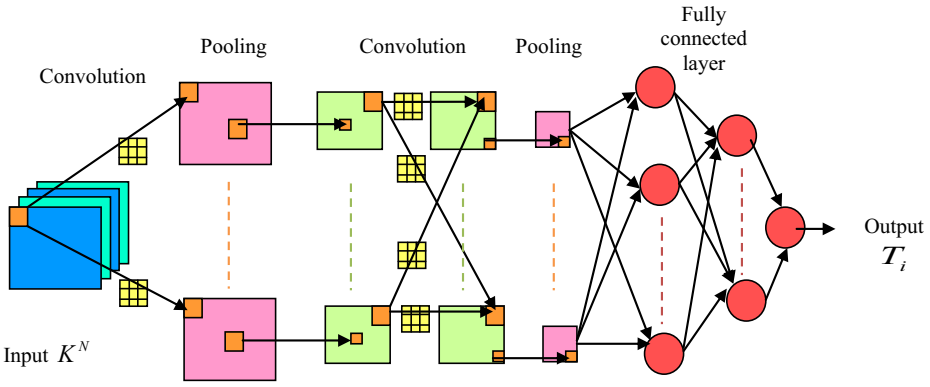


**Figure 2. The architecture of Deep Q-network**

the DHO algorithm include the position of the deer and the wind angle, both of which play pivotal roles in successful hunting.

One essential aspect that greatly impacts the hunting outcome is cooperation among the hunters. Just like how cooperation among real-life hunters can lead to improved hunting efficiency and better results, the DHO algorithm leverages cooperation among its simulated hunters to enhance its optimization process.

The algorithmic procedure of the DHO algorithm can be summarized as follows: elaborate on the specific steps and mathematical formulations involved in the DHO algorithm procedure.

By mimicking the cooperative hunting strategies employed by real-life hunters, the DHO algorithm effectively explores the solution space and finds optimal solutions for complex optimization problems. Its ability to adapt and cooperate makes it a versatile and effective optimization tool across various domains.

***(i) Initialization***

The initialization process sets the starting positions of the hunters, assigns initial wind angles, and defines the level of cooperation.

$$X = \{X_1, X_2, \ldots\ldots, X_c\}; \qquad 1 < k \leq c \tag{9}$$

where $c$ denotes the total count of hunters, which is the solution in the population $X$.

***(ii) Evaluate fitness function***

The fitness function is utilized to maximize the heart disease detection rate in the optimization process.

$$\varpi \frac{1}{c_t} \sum_{t=1}^{c_t} [R_t - T_i]^2 \tag{10}$$

where $c_t$ denotes the total count of samples, $R_t$ is an estimated or targeted output and is the result achieved through the Deep Q-network classifier.

***(iii) Parametric initialization***

The deer's position angle and the wind angle are crucial parameters in determining the optimal positions of hunters. Assuming the search space is circular, the wind angle represents the circumference of the circle.

$$\varphi_t = 2\pi q \tag{11}$$

where $q$ specifies the random number lies in the range of [0, 1] and signifies the current iteration. However, the position angle of the deer is computed as,

$$\psi_t = \phi + \pi \tag{12}$$

where $\phi$ denotes the wind angle.

***(iv) Update the position***

Initially, the location of the best space is not known and the algorithm assumes a candidate solution nearer to the optimal value that is estimated according to the fitness parameter as the optimal solution. Let us represent two solutions, such as the leader position denoted as $X^{lead}$, which is the initial best location of the hunter, and the successor location $X^{successor}$, which is the location of the succeeding hunter.

***Case 1: Propagation through a leader's position***

Once the best position is defined, each individual tries to achieve the optimal position and hence the mechanism of updating the location starts. Therefore, the encircling behavior is computed as follows:

$$X_{t+1} = X^{lead} - Y.s.\left|V \times X^{lead} - X_t\right| \tag{13}$$

where $X_t$ is the position at the current iteration, while $X_{t+1}$ is the location at the next iteration. $Y$ and $V$ denote the coefficient vectors and $s$is a random number established by assuming wind speed ranges from 0 to 2.

Assuming that $X^{lead} > X_t$,

$$X_{t+1} = X^{lead} - Y.s\left|V \times X^{lead} - X_t\right| \tag{14}$$

$$X_{t+1} = X^{lead} - Y.s.V\,X^{lead} + Y.s.X_t \tag{15}$$

$$X_{t+1} = X^{lead}(1 - Y.s.V) + Y.s.X_t \tag{16}$$

From PO,

$$X_{t+1} = l^* + Z(l^* - X_t) \tag{17}$$

$$X_{t+1} = l^* + Z\,l^* - Z\,X_t \tag{18}$$

$$l^* = \frac{X_{t+1} + Z\,X_t}{1 + Z} \tag{19}$$

As $l^*$ holds the winner of the constituency, it can be substituted for the leader's position of DHO.

$$X_{t+1} = \frac{X_{t+1} + Z\,X_t}{1 + Z}\,(1 - Y.s.V) + Y.s.X_t \tag{20}$$

$$X_{t+1} = \frac{X_{t+1}}{1 + Z}(1 - Y.s.V) + \frac{Z}{1 + Z}\,X_t\,(1 - Y.s.V) + Y.s.X_t \tag{21}$$

$$X_{t+1} - \frac{X_{t+1}}{1 + Z}\,(1 - Y.s.V) = \frac{Z}{1 + Z}\,X_t\,(1 - Y.s.V) + Y.s.X_t \tag{22}$$

$$X_{t+1}\left(1 - \frac{1 - Y.s.V}{1 + Z}\right) = \frac{Z}{1 + Z}\,X_t\,(1 - Y.s.V) + Y.s.X_t \tag{23}$$

$$X_{t+1}\left(\frac{1 + Z - 1 + Y.s.V}{1 + Z}\right) = \frac{Z}{1 + Z}\,X_t\,(1 - Y.s.V) + Y.s.X_t \tag{24}$$

$$X_{t+1} = \frac{1 + Z}{Z + Y.s.V}\left[\frac{Z}{1 + Z}\,X_t\,(1 - Y.s.V) + Y.s.X_t\right] \tag{25}$$

where $Z$ lies in the range of [0, 1]. The coefficient vectors are computed as follows:

$$Y = \frac{1}{4}\log\left(t + \frac{1}{t_{\max}}\right) U \tag{26}$$

$$V = 2.C \tag{27}$$

Here, $t_{\max}$ is a maximum number of iteration, which is a parameter that lies in the limit of and $C$ is a random value between the interval of $[0, 1]$.

***Case 2: Propagation through position angle***

To improve search space, the principle is prolonged by assuming the position angle in the update rule. The angle calculation is significant to estimate the location of a hunter. The angle of visualization of prey is modeled as,

$$E_t = \frac{\pi}{8} \times q \tag{28}$$

According to the change between the visual angle of the prey and the wind angle, the updated equation of the position angle is expressed as,

$$W_t = \phi_t - E_t \tag{29}$$

where $\phi$ is the wind angle and position angle is updated to the next iteration, which is expressed as follows:

$$\psi_{t+1} = \psi_t + W_t \tag{30}$$

The position update is represented as,

$$X_{t+1} = X^{lead} - s.\left|\mathrm{Cos}\,(\delta) \times X^{lead} - X_t\right| \tag{31}$$

***Case 3: Propagation through successor***

The value of the vector is considered to be less than 1. Hence, the position update is according to the successor position instead of the initial optimal solution obtained.

$$X_{t+1} = X^{sucessor} - Y.s.\left|V \times X^{successor} - X_t\right| \tag{32}$$

***(v) Termination***

The position update is continued at each iteration until the optimal position is satisfied. Algorithm 1 portrays the pseudo-code of the PDHO algorithm.

**Algorithm 1.** Pseudo code of PDHO

| Sl. No | Pseudo Code of PDHO |
|---|---|
| 1 | Input: Initialize population $X$ |
| 2 | Output: First best solution $X^{leda}$ and second best solution $X^{successor}$ |
| 3 | Begin |
| 4 | While $(t < t_{\max})$ |
| 5 | For each solution in the population |
| 6 | Determine the fitness function |
| 7 | Update $E$, $W$, $s$, $Y$, $V$, and $U$ |
| 8 | if$(s < 1)$ |
| 9 | if $(\lvert V\rvert 1)$ |
| 10 | Update the position of the individual using Eq. (25) |
| 11 | else |
| 12 | Update the position of an individual based on Eq. (31) |
| 13 | end if |
| 14 | else |
| 15 | Update the individual's position according to Eq. (32) |
| 16 | end if |
| 17 | end for |
| 18 | Determine the fitness of each solution |
| 19 | Update $X^{lead}$ and $X^{successor}$ |
| 20 | $t = t + 1$ |
| 21 | end while |
| 22 | Return $X^{lead}$ |
| 23 | Terminate |

## 4. Results and Discussion

This section presents the results and discussion of the developed PDHO-based Deep Q-network, focusing on the evaluation metrics for heart disease detection.

### 4.1. Experimental setup

The implementation of the developed PDHO-based Deep Q-network is carried out using the PYTHON tool on a system with 4GB RAM and an Intel Core-i3 processor. The dataset utilized in this research is the gene expression omnibus dataset [GSE98583] specified in GEO (2021).

### 4.2. Dataset description

Gene expression omnibus is a dataset for gene expression profiling and RNA methylation profiling maintained using National Center for Biotechnology Information. Such accurate-throughput screening genomics information is obtained from microarray experimental information and this information needs to be converted into minimum information about a microarray experiment.

A total of 12 non-diabetic male patients with stable CAD from the North Indian population were included. They were divided into two groups: six with single-vessel disease and six with triple-vessel disease based on coronary angiogram findings. Lesion severity was assessed using a modified Gensini Scoring system, and a control group of individuals with atypical angina and normal coronary angiograms was recruited for comparison (GEO, 2021).

### 4.3. Evaluation metrics

The performance of the developed PDHO-based Deep Q-network is analyzed by considering key evaluation metrics, including accuracy, sensitivity (recall), and specificity. These metrics provide valuable insights into the model's overall accuracy, its ability to correctly identify positive cases (heart disease) as true positives (sensitivity/recall), and its capability to correctly identify negative cases (healthy) as true negatives (specificity). By examining these metrics, we gain a comprehensive understanding of the model's effectiveness and its potential for accurate heart disease detection.

***(i) Accuracy:*** Accuracy is defined as a degree of closest value of measurements of a quantity. In other words, accuracy is also termed as the standard of the correct or precise measurement.

$$Accuracy = \frac{TP + TN}{(TP + TN + FP + FN)} \tag{33}$$

where $FP$ and $FN$ represent the count of false positives and false negatives, respectively. The number of true negatives is termed as $TN$.
***(ii) Sensitivity:*** Sensitivity, known as the true positive rate, measures the proportion of positive cases correctly identified by a diagnostic test or model, indicating how well it captures true positive instances.

$$Sensitivity = \frac{TP}{TP + FN} \tag{34}$$

where $TP$ denotes the number of true positives and the count of false negatives is represented as $FN$.
***(iii) Specificity:*** Specificity, termed as the true negative rate, measures the proportion of true negative cases correctly identified by a diagnostic test or model, indicating its effectiveness in correctly classifying negative instances (e.g., healthy cases) among all the actual negatives in the dataset.

$$Specificity = \frac{TN}{TN + FP} \tag{35}$$

## 4.4. Performance analysis

In this section, we present the performance analysis of the developed PDHO-based Deep Q-network in terms of evaluation metrics.

*4.4.1. Analysis based on training data*

Figure 3 illustrates the performance analysis of the developed PDHO-based Deep Q-network in terms of evaluation metrics, namely accuracy, sensitivity, and specificity, while varying the training data percentage.

(a) Analysis of Accuracy:

When the training data are set to 60%, the PDHO-based Deep Q-network achieved accuracy of 0.732 with 5 iterations, 0.742 with 10 iterations, 0.752 with 15 iterations, and 0.759 with 20 iterations. When the training data are increased to 90%, the accuracy improved to 0.925 with 5 iterations, 0.934 with 10 iterations, 0.941 with 15 iterations, and 0.950 with 20 iterations.

(b) Analysis of Sensitivity:

With 60% training data, the PDHO-based Deep Q-network achieved sensitivity of 0.784 with 5 iterations, 0.792 with 10 iterations, 0.798 with 15 iterations, and 0.809 with 20 iterations. When the training data were raised to 90%, sensitivity increased

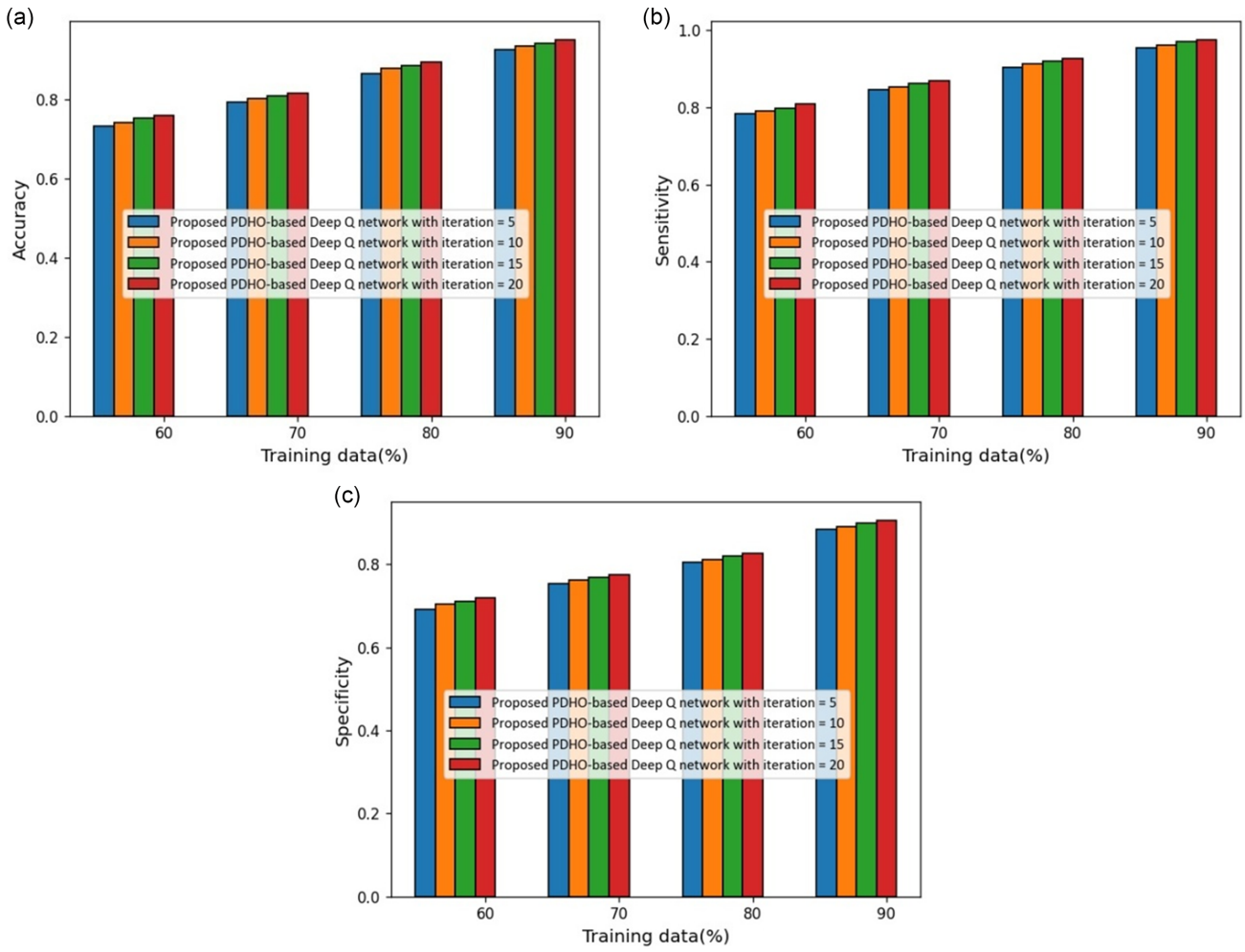


**Figure 3. Performance analysis using training data (a) accuracy, (b) sensitivity, and (c) specificity**

to 0.954 with 5 iterations, 0.962 with 10 iterations, 0.969 with 15 iterations, and 0.974 with 20 iterations.

(c) Analysis of Specificity:

For 60% training data, the PDHO-based Deep Q-network achieved specificity of 0.691 with 5 iterations, 0.705 with 10 iterations, 0.711 with 15 iterations, and 0.719 with 20 iterations. With 90% training data, the specificity improved to 0.884 with 5 iterations, 0.892 with 10 iterations, 0.899 with 15 iterations, and 0.905 with 20 iterations.

These results demonstrate how the PDHO-based Deep Q-network's performance in terms of accuracy, sensitivity, and specificity varies with different amounts of training data, providing valuable insights into the model's effectiveness and its ability to adapt to varying dataset sizes during heart disease detection.

*4.4.2. Analysis based on the number of features*

Figure 4 presents the analysis of the developed PDHO-based Deep Q-network based on different numbers of features.

(a) Accuracy Analysis:

With 2000 features, the PDHO-based Deep Q-network achieved accuracy of 0.851 with 5 iterations, 0.863 with 10 iterations, 0.875 with 15 iterations, and 0.884 with 20 iterations. With 8000 features, the accuracy improved to 0.883 with 5 iterations, 0.890 with 10 iterations, 0.894 with 15 iterations, and 0.897 with 20 iterations.

(b) Sensitivity Analysis:

For 2000 features, the PDHO-based Deep Q-network achieved sensitivity of 0.891 with 5 iterations, 0.905 with 10 iterations, 0.912 with 15 iterations, and 0.920 with 20 iterations. With 8000 features, the sensitivity improved to 0.925 with 5 iterations, 0.934 with 10 iterations, 0.943 with 15 iterations, and 0.951 with 20 iterations.

(c) Specificity Analysis:

With 2000 features, the PDHO-based Deep Q-network achieved specificity of 0.792 with 5 iterations, 0.799 with 10 iterations, 0.806 with 15 iterations, and 0.812 with 20 iterations. With 8000 features, the specificity improved to 0.819 with

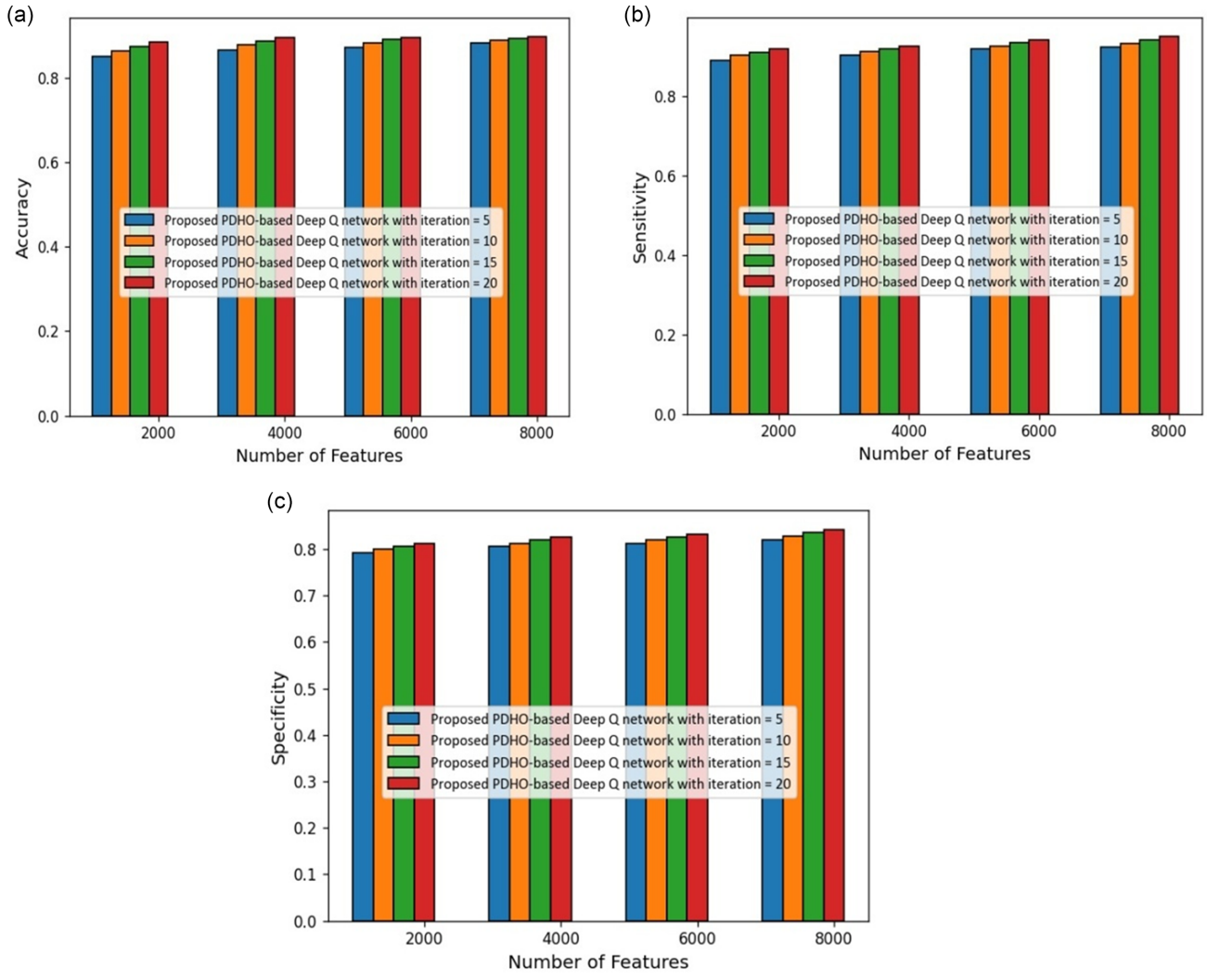


**Figure 4. Performance analysis using several features (a) accuracy, (b) sensitivity, and (c) specificity**

5 iterations, 0.828 with 10 iterations, 0.836 with 15 iterations, and 0.841 with 20 iterations.

These results demonstrate the performance of the PDHO-based Deep Q-network in terms of accuracy, sensitivity, and specificity, varying with the number of features used. The analysis provides insights into how different feature sets influence the model's effectiveness in detecting heart disease.

## 4.5. Comparative analysis

The performance of the developed PDHO-based Deep Q-network is analyzed through a comparison with several conventional approaches, including U-net CNN (Xiao et al., 2020), DNN (Ashraf et al., 2019), Deep CNN (Mehmood et al., 2021), and FRDLS (Wang et al., 2020). This comparative analysis allows us to evaluate the effectiveness and superiority of the proposed PDHO-based Deep Q-network in heart disease detection, providing valuable insights into its performance and potential advantages over existing methods.

Figure 5 presents the comparative analysis of the developed PDHO-based Deep Q-network and conventional approaches in terms of evaluation metrics.

(a) Accuracy Analysis:

At training data = 60%, the PDHO-based Deep Q-network achieved an accuracy of 0.742, outperforming U-net CNN (0.627), DNN (0.660), Deep CNN (0.688), and FRDLS (0.704). At training data = 90%, the PDHO-based Deep Q-network achieved an accuracy of 0.934, surpassing U-net CNN (0.751), DNN (0.777), Deep CNN (0.806), and FRDLS (0.851).

(b) Sensitivity Analysis:

At training data = 60%, the sensitivity of the PDHO-based Deep Q-network was 0.792, outperforming U-net CNN (0.697), DNN (0.700), Deep CNN (0.710), and FRDLS (0.740). At training data = 90%, the sensitivity of the PDHO-based Deep Q-network was 0.962, surpassing U-net CNN (0.780), DNN (0.797), Deep CNN (0.828), and FRDLS (0.887).

(c) Specificity Analysis:

At training data = 60%, the specificity of the PDHO-based Deep Q-network was 0.705, outperforming U-net CNN (0.605), DNN (0.624), Deep CNN (0.650), and FRDLS (0.676). At training data = 90%, the specificity of the PDHO-based Deep

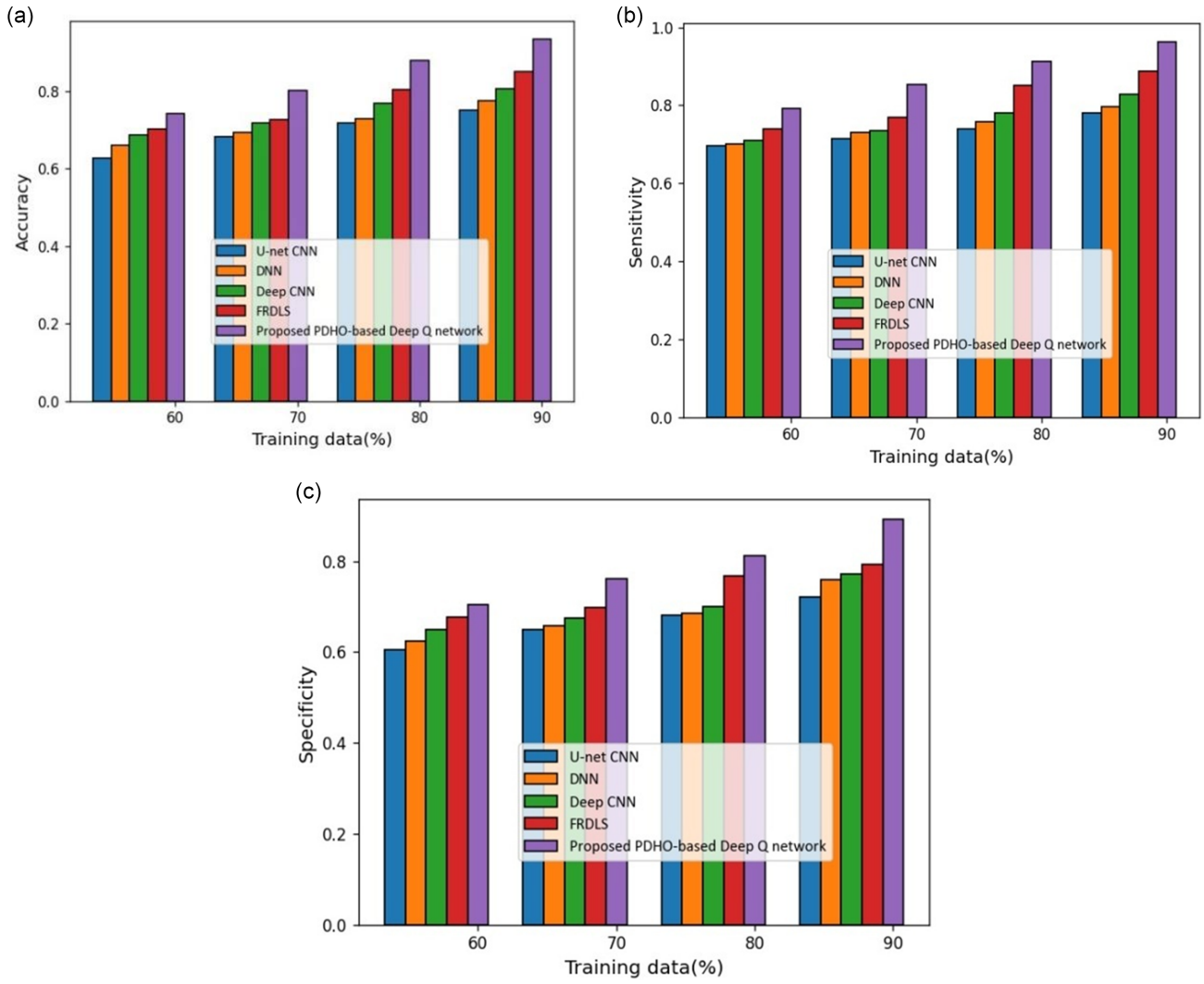


**Figure 5. Comparative analysis based on training data (a) accuracy, (b) sensitivity, and (c) specificity**

Q-network was 0.892, surpassing U-net CNN (0.721), DNN (0.758), Deep CNN (0.773), and FRDLS (0.793).

The analysis demonstrates that the developed PDHO-based Deep Q-network exhibits superior performance in terms of accuracy, sensitivity, and specificity compared to the conventional approaches across different training data percentages.

### 4.6. Statistical tests

Table 1 presents a statistical tests of the developed PDHO-based Deep Q-network and existing methods. At a training data percentage of 90%, the PDHO-based Deep Q-network outperformed the conventional approaches with an accuracy of 0.934, sensitivity of 0.962, and specificity of 0.892. The existing methods achieved accuracies of 0.751 (U-net CNN), 0.777 (DNN), 0.806 (Deep CNN), and 0.851 (FRDLS) at the same training data percentage. Additionally, at a training data percentage of 60%, the specificity of the conventional methods (U-net CNN, DNN, Deep CNN, and FRDLS) was 0.605, 0.624, 0.650, and 0.676, respectively. The table highlights the superior performance of the proposed PDHO-based Deep Q-network, with maximum accuracy, sensitivity, and specificity.

**Table 1. Statistical tests of PDHO-based Deep Q-network and existing methods where training data = 90%**

| | Accuracy | Sensitivity | Specificity |
|---|---|---|---|
| **U-net CNN** | 75.1% | 78.0% | 72.11% |
| **DNN** | 77.7% | 79.7% | 75.8% |
| **Deep-CNN** | 80.6% | 82.8% | 77.3% |
| **FRDLS** | 85.1% | 88.7% | 79.3% |
| **Proposed PDHO-based Deep Q-network** | **93.4%** | **96.2%** | **89.2%** |

## 5. Conclusion

Heart disease is a significant global health concern, leading to high mortality rates, especially in Western countries. Early detection of heart disease is crucial for better medical intervention and reducing patient mortality. While various deep learning-based techniques have been developed for heart disease detection, they often face challenges with imbalanced data and inadequate feature representation. To address these issues, a novel approach called PDHO-based Deep Q-network is proposed. This method utilizes gene expression data as input, undergoes pre-processing via Box–Cox transformation, and then fuses features using a correlation-based Deep Q-network. The heart disease detection is performed using a Deep Q-network classifier, trained using the PDHO algorithm. The proposed PDHO-based Deep Q-network achieved remarkable accuracy 93.4%, sensitivity 96.2%, and specificity 89.2% in heart disease detection. Future work may involve integrating other optimization algorithms to further enhance accuracy levels. Detecting heart disease at an early stage using advanced techniques holds the promise of significantly improving patient outcomes and reducing mortality rates worldwide.

## Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

## Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

## Data Availability Statement

The data that support the findings of this study are openly available in NCBI at https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98583.

## References

Abdeltawab, H., Khalifa, F., Taher, F., Alghamdi, N. S., Ghazal, M., Beache, G., . . . , & El-Baz, A. (2020). A deep learning-based approach for automatic segmentation and quantification of the left ventricle from cardiac cine MR images. *Computerized Medical Imaging and Graphics*, *81*, 101717. https://doi.org/10.1016/j.compmedimag.2020.101717

Aher, C. N., & Jena, A. K. (2023a). Heart disease detection from gene expression data using optimization driven deep Q-network. In *International Conference on Frontiers of Intelligent Computing: Theory and Applications*, 327, 601–611. https://doi.org/10.1007/978-981-19-7524-0_53

Aher, C. N., & Jena, A. K. (2023b). Improved invasive weed bird swarm optimization algorithm (IWBSOA) enabled hybrid deep learning classifier for diabetic prediction. *Journal of Ambient Intelligence and Humanized Computing*, *14*(4), 3929–3945. https://doi.org/10.1007/s12652-022-04462-z

Ali, F., El-Sappagh, S., Islam, S. R., Kwak, D., Ali, A., Imran, M., & Kwak, K. S. (2020). A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion. *Information Fusion*, *63*, 208–222. https://doi.org/10.1016/j.inffus.2020.06.008

Ashraf, M., Rizvi, M. A., & Sharma, H. (2019). Improved heart disease prediction using deep neural network. *Asian Journal of Computer Science and Technology*, *8*(2), 49–54. https://doi.org/10.51983/ajcst-2019.8.2.2141

Askari, Q., Younas, I., & Saeed, M. (2020). Political optimizer: A novel socio-inspired meta-heuristic for global optimization. *Knowledge-Based Systems*, *195*, 105709. https://doi.org/10.1016/j.knosys.2020.105709

Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In I. Cohen, Y. Huang, J. Chen, & J. Benesty (Eds.), *Noise reduction in speech processing* (pp. 1–4), Springer. https://doi.org/10.1007/978-3-642-00296-0_5

Brammya, G., Praveena, S., Ninu Preetha, N. S., Ramya, R., Rajakumar, B. R., & Binu, D. (2019). Deer hunting optimization algorithm: A new nature-inspired meta-heuristic paradigm. *The Computer Journal, bxy133*. https://doi.org/10.1093/comjnl/bxy133

Candemir, S., White, R. D., Demirer, M., Gupta, V., Bigelow, M. T., Prevedello, L. M., & Erdal, B. S. (2020). Automated coronary artery atherosclerosis detection and weakly supervised localization on coronary CT angiography with a deep 3-dimensional convolutional neural network. *Computerized Medical Imaging and Graphics*, *83*, 101721. https://doi.org/10.1016/j.compmedimag.2020.101721

Chen, A. H., Huang, S. Y., Hong, P. S., Cheng, C. H., & Lin, E. J. (2011). HDPS: Heart disease prediction system. In *Proceedings of 2011 Computing in Cardiology*, 557–560. Retrieved from: https://ieeexplore.ieee.org/abstract/document/6164626/metrics#metrics

Guo, K., Mei, H., Xie, X., & Xu, X. (2019). A convolutional neural network feature fusion framework with ensemble learning for EEG-based emotion classification. In *2019 IEEE MTT-S International Microwave Biomedical Conference, 1*, 1–4. Retrieved from: https://ieeexplore.ieee.org/abstract/document/8777738

Han, S. S., Kim, M. S., Lim, W., Park, G. H., Park, I., & Chang, S. E. (2018). Classification of the clinical images for benign and malignant cutaneous tumors using a deep learning algorithm. *Journal of Investigative Dermatology*, *138*(7), 1529–1538. https://doi.org/10.1016/j.jid.2018.01.028

Khan, R. (2020). Artificial bee colony-based general adversarial network for liver cancer detection using CT images. *Multimedia Research*, *3*(4), 1–11. https://doi.org/10.46253/j.mr.v3i4.a1

Li, H., Wang, X., Liu, C., Zeng, Q., Zheng, Y., Chu, X., . . . , & Karmakar, C. (2020). A fusion framework based on multi-domain features and deep learning features of phonocardiogram for coronary artery disease detection. *Computers in Biology and Medicine*, *120*, 103733. https://doi.org/10.1016/j.compbiomed.2020.103733

Maciejewski, R., Pattath, A., Ko, S., Hafen, R., Cleveland, W. S., & Ebert, D. S. (2012). Automated box-cox transformations for improved visual encoding. *IEEE Transactions on Visualization and Computer Graphics*, *19*(1), 130–140. https://doi.org/10.1109/TVCG.2012.64

Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of heart disease using deep convolutional neural networks. *Arabian Journal for Science and Engineering*, *46*, 3409–3422. https://doi.org/10.1007/s13369-020-05105-1

Neelima, E., & Babu, M. P. (2020). Genome feature optimization and coronary artery disease prediction using cuckoo search. *Computer Science and Information Technologies*, *1*(3), 106–115. https://doi.org/10.11591/csit.v1i3.p106-115

Nikdelfaz, O., & Jalili, S. (2018). Disease genes prediction by HMM based PU-learning using gene expression profiles. *Journal of Biomedical Informatics*, *81*, 102–111. https://doi.org/10.1016/j.jbi.2018.03.006

Padmanabhan, V., & Semmlow, J. L. (1994). Dynamical analysis of diastolic heart sounds associated with coronary artery disease. *Annals of Biomedical Engineering*, *22*, 264–271. https://doi.org/10.1007/BF02368233

Rairikar, A., Kulkarni, V., Sabale, V., Kale, H., & Lamgunde, A. (2017). Heart disease prediction using data mining techniques. In *2017 International Conference on Intelligent Computing and Control*, 1–8. https://doi.org/10.1109/I2C2.2017.8321771

Sasaki, H., Horiuchi, T., & Kato, S. (2017). A study on vision-based mobile robot learning by deep Q-network. In *56th Annual Conference of the Society of Instrument and Control Engineers of Japan*, 799–804. https://doi.org/10.23919/SICE.2017.8105597

Semmlow, J., & Rahalkar, K. (2007). Acoustic detection of coronary artery disease. *Annual Review of Biomedical Engineering*, *9*, 449–469. https://doi.org/10.1146/annurev.bioeng.9.060906.151840

Sharma, S., & Parmar, M. (2020). Heart diseases prediction using deep learning neural network model. *International Journal of Innovative Technology and Exploring Engineering*, *9*(3), 2244–2248. https://doi.org/10.35940/ijitee.C9009.019320

The National Center for Biotechnology Information. *Global gene expression studies for coronary artery disease and its severity in North Indian population*. Retrieved from: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE98583

Wang, Z., Zhu, Y., Li, D., Yin, Y., & Zhang, J. (2020). Feature rearrangement based deep learning system for predicting heart failure mortality. *Computer Methods and Programs in Biomedicine*, *191*, 105383. https://doi.org/10.1016/j.cmpb.2020.105383

Xiao, C., Li, Y., & Jiang, Y. (2020). Heart coronary artery segmentation and disease risk warning based on a deep learning algorithm. *IEEE Access*, *8*, 140108–140121. https://doi.org/10.1109/ACCESS.2020.3010800

**How to Cite:** Aher, C. N. (2023). Enhancing Heart Disease Detection Using Political Deer Hunting Optimization-Based Deep Q-Network with High Accuracy and Sensitivity. *Medinformatics*. https://doi.org/10.47852/bonviewMEDIN32021377