Enhancing Cardiovascular Disease Prediction via Hybrid Deep Learning Architectures: A Step Towards Smart Healthcare

1st Aitzaz Ahmed Murtaza
Dept. of Mechanical Engineering
CUST
Islamabad, Pakistan
aitzaz332@gmail.com

2nd Amina Saher
School of Natural Sciences
NUST
Islamabad, Pakistan
amnasaher332@gmail.com

3rd Hassan Mohyuddin
Dept. of Electrical Engineering
CUST
Islamabad, Pakistan
mee213003@cust.pk

4th Syed Kumayl Raza Moosavi
SEECS
NUST
Islamabad, Pakistan
smoosavi.msai21seecs@seecs.edu.pk

5th Muhammad Hamza Zafar
Dept. of Engineering Sciences
University of Agder
Grimstaad, Norway
muhammad.h.zafar@uia.no

6th Filippo Sanfilippo
Dept. of Engineering Sciences
University of Agder
Grimstaad, Norway
filippo.sanfilippo@uia.no

***Abstract*—Cardiovascular disease presents a serious and increasing global health challenge, making a substantial contribution to morbidity and mortality rates on a global scale. This research study presents a novel methodology for predicting Cardiovascular Diseases by employing a, recently developed, metaheuristic optimisation algorithm within a neural network framework. The Coati Optimisation Algorithm (COA) is employed in an artificial neural network (ANN) to enhance the predictive accuracy of outcomes related to Cardiovascular Diseases. The enhanced performance of the COA can be ascribed to its adept utilisation of both exploration and exploitation phenomena. This research employs publicly available datasets pertaining to heart and stroke disorders, integrating two datasets focused on heart disease and one dataset focused on stroke disease. A comparison analysis is undertaken between the proposed COA-ANN and existing approaches, namely Particle Swarm Optimizer based ANN (PSO-ANN), Grey Wolf Optimizer based ANN (GWO-ANN), and backpropagation based ANN (BP-ANN). The findings of the study indicate that the COA-ANN model exhibits the highest level of predictive accuracy. The COA-ANN outperformed the other three networks, namely GWO-ANN, PSO-ANN, and BP-ANN, with an average accuracy of 98.43%. As a result, the utilisation of the COA-ANN leads to an improvement in predictive accuracy for these datasets, with an increase of up to 2.64%. Additional assessment metrics, such as F1-Score, Precision, and Recall, provide more insight into the balanced performance of the COA-ANN architecture when applied to imbalanced class datasets. These results prove that the integration of nature-inspired algorithms with cardiovascular diseases (CVDs) is a promising direction for future research.**

***Index Terms*—Metaheuristic Algorithms, Cardiovascular Diseases (CVDs), Machine Learning, Artificial Neural Network**

I. Introduction

Cardiovascular diseases (CVDs) are a major cause of mortality in both developed and developing countries. Significant data is generated during the diagnosis and treatment of CVDs. The application of machine learning (ML) techniques enables the efficient detection of CVDs through the analysis and evaluation of data. CVDs are the leading cause of human mortality, resulting in 17.9 million deaths annually, according to the World Health Organisation (WHO) [1].

CVDs, including heart disease and stroke, have many risk factors. Several variables may cause certain health issues. These include smoking, gender, bad lifestyle choices, family history, lack of exercise, hypertension, obesity, and excessive alcohol intake. Genetic factors including diabetes and hypertension are substantial risk factors for stroke and heart diseases [2]. Heart disease causes chest pain, palpitations, irregular heartbeats, dyspnea, and weariness. Therefore, timely identification of cardiovascular diseases is essential to reduce the risk of mortality.

Similarly, stroke disease is the second leading cause of human mortality, accounting for approximately 11% of documented global deaths [3]. Around 5% of individuals aged 14 and above may encounter the consequences of a stroke. There is significant variation in stroke mortality rates among countries, with higher death rates observed in countries with lower average wages and lower death rates observed in countries with higher average wages [4]. The risk factors for stroke are similar to those for heart disease. Stroke is a pathological condition that requires immediate medical attention and intervention. Therefore, promptly recognising a stroke can lead to a reduction in mortality rates linked to this ailment.

Hospitals generate significant volumes of data when diagnosing life-threatening illnesses. The collected data is currently being utilised in the field of artificial intelligence (AI) to develop novel methodologies for cardiovascular disease prediction. AI has become an important factor in cardiology due to advancements in data-related technologies [5]. The researchers utilised various data mining and preprocessing techniques to effectively input the data into ML models for accurate and efficient prediction of CVDs. Various ML techniques are currently utilised for predicting CVDs.

In the field of ML, it is customary to divide a large dataset into smaller subsets. The subsets are chosen meticulously to include only the most pertinent features, which display notable patterns identified through ML models [6]. These patterns are essential for facilitating decision-making. Various ML techniques are frequently used for predicting CVDs. The algorithm used in this study is a hybrid deep learning (DL) architecture utilizing the Artificial Neural Network (ANN). ML methodologies can be categorised into three main groups: supervised learning, unsupervised learning, and reinforcement learning [7].

This study employs three open-source datasets and applies various ML techniques for classification and optimization. The contributions are elaborated as follows:

- Prediction and classification of three CVD-based datasets are displayed using a novel metaheuristic DL technique i.e., COA-ANN.
- A comparative study is conducted with two metaheuristic-based NN models, namely: Particle Swarm Optimisation (PSO), Grey-Wolf Optimizer (GWO), and the traditional backpropagation based ANN architecture
- The findings of the study suggest the COA-ANN model outperforms traditional backpropagation techniques and competitive metaheuristic algorithms.

II. Related Work

In the realm of deep learning (DL), metaheuristic algorithms have been widely utilized as a powerful optimization tool specially in cases where there are class imbalances in the dataset leading to local and global minima solutions. Also, metaheuristic algorithms are becoming more critical in computational intelligence because they are flexible, adaptive, and have an extensive search capacity. From an inspirational standpoint, these algorithms can be subdivided into three major categories; biology based - inspired by the joint movement of mammals and animals for hunting - evolutionary based - inspired by the process of biological evolution - and human interaction based - inspired by human behaviors, interactions, and decision-making processes [8].

In terms of number of solutions, these algorithms can be classified into two categories. The first category concerns trajectory-based algorithms, which deal with a single solution and its improvement as the algorithm moves through the search space for finding the optimum solution. Trajectory-based algorithms include Simulated Annealing (SA), Tabu Search (TS), Greedy Randomized Adaptive Search Procedure (GRASP) and Variable Neighborhood Search (VNS) [9]. Whereas, the second category is about population based algorithms, which deal with a set of solutions and keep on improving them based on the information obtained from each other. Some examples include Genetic Algorithm (GA) [10], Grey Wolf Optimizer (GWO) [11], Particle Swarm Optimizer (PSO) [12], Ant Colony Optimizer (ACO) [13], and Red Deer Algorithm (RDA) [14]. Similarly, Coati optimization algorithm (COA) [15], a population based algorithm, is designed on the natural behavior of coatis. This study employs the COA as an optimizer for a DL based ANN architecture to find the best solution to minimize the cost. COA is chosen because of its capabilities of finding the best cost and avoiding local minima. Another feature of COA is its capability to effectively utilize both exploration and exploitation for finding the best cost value and better convergence.

III. Proposed Technique

A. Coati Optimization Algorithm

COA mimics the strategies adopted by coatis while hunting iguanas, and the behavior when confronting and fleeing from the predator [15]. These strategies are an intelligent phenomenon that is the base of developing COA. The mathematical model of COA is devised as under.

Mathematical Model: The initialization of COA begins with the initialization of positions of coatis randomly, using the following equation;

$$X_i : x_{i.j} = lb_j + r.\left(ub_j - lb_j\right), \tag{1}$$

where $i = 1, 2, 3, \ldots, N$, $j = 1, 2, 3, \ldots, M$, X_i is the position of i^{th} coati, number of coatis are N and number of decision variables are m. ub_j and lb_j are the upper and lower bound of j^{th} decision variable respectively. r is the random number which is a real number from 0 to 1.

Population of coatis in COA is represented by the following matrix:

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} x_{11} & \ldots & x_{1j} & \ldots & x_{1m} \\ \vdots & \ddots & \vdots & . & \vdots \\ x_{i1} & \ldots & x_{ij} & \ldots & x_{im} \\ \vdots & . & \vdots & \ddots & \vdots \\ x_{N1} & \ldots & x_{Nj} & \ldots & x_{Nm} \end{bmatrix} \tag{2}$$

The objective function or fitness function is evaluated against each position of coatis and is given by the following equation:

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N*1} = \begin{bmatrix} F\left(X_1\right) \\ \vdots \\ F\left(X_i\right) \\ \vdots \\ F\left(X_N\right) \end{bmatrix}_{N*1}, \tag{3}$$

where, F is the objective function vector which is obtained during the propagation of algorithm against each position.

The movement of coatis is based on the two strategies; hunting the prey and escaping the predator.

Strategy 1: Hunting the iguana (exploration):

Coatis usually look for the iguanas climbed on the trees. One of the two groups formed, scares the prey by climbing and other waits on the ground for attacking and hunting, as

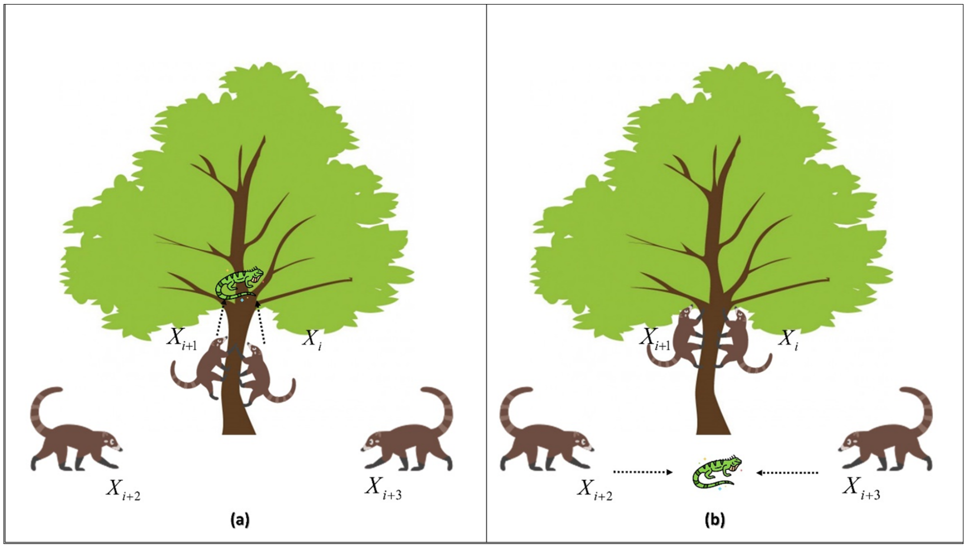

Fig. 1. Hunting strategy of COA. (a) Half of the coatis are attacking the Iguana. (b) The other half hunting the Iguana once fell on the ground.

illustrated in Fig. 1. This strategy explains the exploration ability of COA in the search space.

The first half of coatis is represented by the following mathematical equation:

$$X_i^{p1}: x_{ij}^{p1} = x_{ij} + r.\left(\text{ Iguana }_j - \text{ I. }_{ij}\right), \tag{4}$$

where $i = 1,2,3,\ldots,\left\lfloor\frac{N}{2}\right\rfloor$ and $j = 1,2,3,\ldots,m$. After the Iguana falls down, the rest of coatis update their position which is simulated by the following equation:

$$\text{Iguana}^G : \text{Iguana}_j^G = lb_i + r\cdot(ub_j - lb_j) \tag{5}$$

$$X_i^{p1} = \begin{cases} x_{ij} + r\cdot\left(\text{ Iguana }_j^G - \text{ I. }_{ij}\right), & F_{\text{Iguana }^G} < F_i \\ x_{ij} + r\cdot\left(x_{ij} - \text{ Iguana }_j^G\right), & \text{else} \end{cases}, \tag{6}$$

where $j = 1,2,3,\ldots$ and $i = \left[\frac{N}{2}\right]+1, \left[\frac{N}{2}\right]+2,\ldots,N$. The condition of updating the position m of coati is given by the following equation:

$$X_i = \begin{cases} X_i^{p1}, & F_i^{p1} < F_i \\ X_i, & \text{else} \end{cases}, \tag{7}$$

whereas X_i^{P1} is the updated position of i^{th} coati, F_i^{P1} is the updated cost function based on the new position. Iguana represents the position of best solution in search space. Iguana^G is the position of prey on the ground and F_{Iguana^G} is its fitness value.

Strategy 2: Escaping the predator (exploitation):

When a predator assaults a coati, it moves to a safer position near its initial position, as illustrated in Fig. 2. This strategy explains the exploitation ability of the COA. For its simulation, the following mathematical equations are used:

$$lb_j^{\text{local}} = \frac{lb_j}{t}, ub_j^{\text{local}} = \frac{ub_j}{t} \tag{8}$$

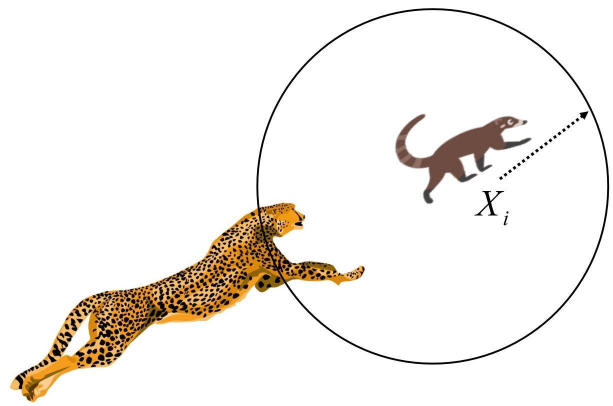

Fig. 2. Escaping the predator.

$$X_i^{p2} = x_{ij} + (1-2r).\left(lb_j^{\text{local}} + r.\left(ub_j^{\text{local}} - lb_j^{\text{local}}\right)\right) \tag{9}$$

This condition for updating the position can be given by the following equation:

$$X_i = \begin{cases} X_i^{p2}, & F_i^{p2} < F_i \\ X_i, & \text{else} \end{cases}, \tag{10}$$

where X_i^{p2} is the new position obtained for i^{th} coati for the second phase of COA, F_i^{p2} is the fitness value of this coati. t is the current iteration number and T is the total iterations. ub_j^{local} and lb_j^{local} are the local upper and lower bound values of the j^th decision variable.

The process from Equation (4) to (10) is repeated until the maximum number of iterations is reached or the convergence value is reached. The program gives output in the form of the best solution obtained so far. This solution is set as the weights and biases of the ANN model to calculate the minimum cost value.

B. Artificial Neural Network (ANN)

Human brain shape and function inspire artificial neural networks (ANNs). They compute complicated mathematical problems to get results. Because of their versatility, back-propagation through learning from data, and capacity to solve complicated problems, ANNs are widely employed. Three layers make up an ANN: input, hidden, and output. Choosing the proper number of hidden layers and neurons is crucial for classification accuracy. If the number of neurons is low, the neural network may not understand all dataset features, reducing accuracy. The neural network would overfit and take a long time if there were too many neurons. So, choosing the proper number of neurons and hidden layers is crucial for neural network accuracy.

Based on the comparison performed on activation functions [16], it is found that the sigmoid function (σ) performs best in the classification problems. Therefore, for this study, the sigmoid function is utilized. Traditionally, the backpropagation algorithm is used in an ANN which is responsible for the adjustment of weights and biases of a neural network [17]. However, they have inherent drawbacks and limitations which include vanishing/exploding gradients, prolonged training times, and poor generalization ability [18]. Therefore, in this study, the authors utilize the optimization ability of metaheuristic algorithms for the optimization of the hyperparameters of the ANN.

IV. Experimental Setup

A. Design of COA based ANN (COA-ANN)

ANN with a single hidden layer is constructed for this study. This ANN consists of three layers; an input layer, a hidden layer, and an output layer as shown in Fig. 3. The hidden layer consists of 10 neurons, and every neuron is connected to all the features of the input layer. The number of input layer neurons depends upon the number of features in the dataset and the number of output layer neurons depend upon the target class of the dataset. The weights and biases of this ANN is updated using COA.

The input vector given as $x_i = [x_i^1, x_i^2, \ldots, x_i^N]$ contains N features which are fed to the input layer of ANN. The corresponding output value is given as $y_i = [y_i]$. The flowchart of using ANN with COA is given in the Fig. 4.

B. Dataset description

Three publicly available datasets were used for this research. Researchers from Bangladeshi hospitals collected the first stroke illness dataset. The dataset contains 5110 patients' data and 10 critical features that help predict disease. The target class indicates if the patient has a stroke disease. The other two datasets are heart disease-related. The ”Framingham” dataset has 4,241 patients aged 32–70. The dataset is collected by researchers for a Framingham, Massachusetts, cardiovascular study. This dataset has 15 characteristics and 16 columns. The third dataset includes Cleveland, Hungary, Switzerland, and VA Long Beach patients. This dataset has 1025 patients, 713 male and 312 female.

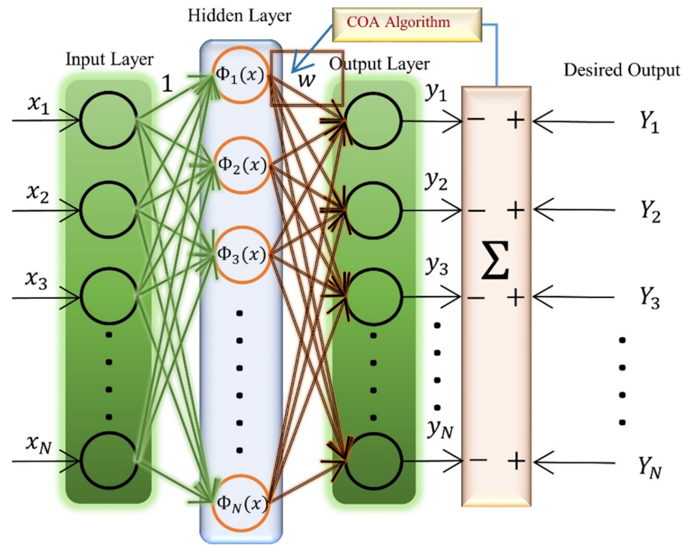

Fig. 3. Structure of COA-ANN.

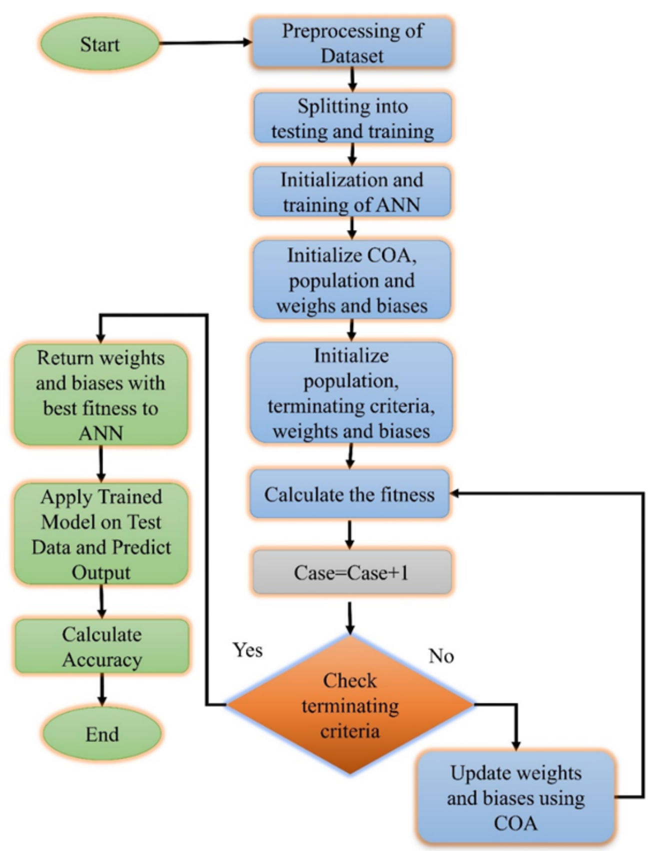

Fig. 4. Flowchart of COA implemented with ANN.

V. Results

The comparative evaluation aimed to discern the efficacy of different approaches: COA-ANN, GWO-ANN, PSO-ANN, and BP-ANN. In the case of Dataset 1, as shown in Table I, COA-ANN showcased remarkable performance with an accuracy of 97.22% and an F1-Score of 97.15%, outpacing other methods. GWO-ANN and PSO-ANN closely followed with competitive results, while BP-ANN displayed comparatively lower accuracy and F1-Score values.

TABLE I
COMPARATIVE ANALYSIS OF EVALUATION MATRICES FOR DATASET 1

Metrics	COA-ANN	GWO-ANN	PSO-ANN	BP-ANN
Accuracy	0.9722	0.9447	0.9624	0.8875
F1-Score	0.9715	0.9358	0.9675	0.8949
Precision	0.9834	0.9470	0.9585	0.8876
Recall	0.9826	0.9334	0.9669	0.8706

Shifting the focus to Dataset 2, as shown in Table II, COA-ANN demonstrated exceptional capabilities by achieving an accuracy of 99.03% and an F1-Score of 98.28%. GWO-ANN also maintained strong performance, while PSO-ANN faced optimization challenges, resulting in lower accuracy and F1-Score values. BP-ANN exhibited the least competitive results. The trends shifted in Dataset 3, as shown in Table III. COA-ANN's accuracy of 99.06% and an impressive F1-Score of 99.54% reflected its profound ability to capture intricate patterns. GWO-ANN performed well, yet PSO-ANN struggled to optimize, leading to modest accuracy and F1-Score figures. BP-ANN's performance remained moderate.

TABLE II
COMPARATIVE ANALYSIS OF EVALUATION MATRICES FOR DATASET 2

Metrics	COA-ANN	GWO-ANN	PSO-ANN	BP-ANN
Accuracy	0.9903	0.9535	0.8032	0.6477
F1-Score	0.9828	0.9576	0.8137	0.6358
Precision	0.9941	0.9482	0.8095	0.6492
Recall	0.9967	0.9528	0.8253	0.6623

TABLE III
COMPARATIVE ANALYSIS OF EVALUATION MATRICES FOR DATASET 3

Metrics	COA-ANN	GWO-ANN	PSO-ANN	BP-ANN
Accuracy	0.9906	0.9754	0.7971	0.7268
F1-Score	0.9954	0.9653	0.7857	0.7126
Precision	0.9832	0.9671	0.7989	0.7365
Recall	0.9981	0.9746	0.7915	0.7204

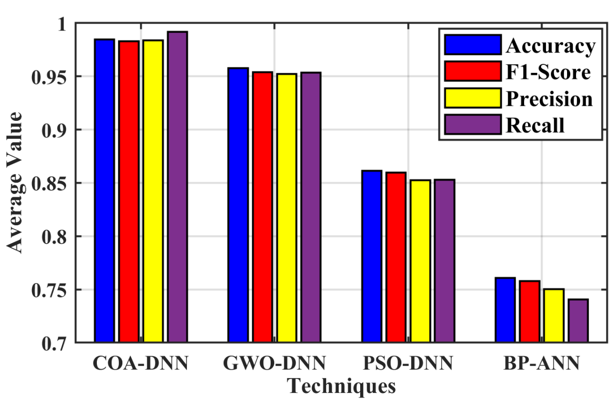

Fig. 5. Average Comparison of Evaluation Matrices.

In general, GWO-ANN upheld its proficiency, while PSO-ANN achieved competitive results. BP-ANN exhibited moderate accuracy and F1-Score figures. The performance variations, overall, across datasets emphasized the influence of data intricacy on technique selection. COA-ANN emerged as a robust performer, while GWO-ANN consistently demonstrated competitiveness. PSO-ANN's results varied, and BP-ANN generally exhibited less competitive outcomes. The optimization methods and neural network architectures significantly impacted the findings. Further investigation, considering additional contextual factors, is recommended to comprehensively understand technique effectiveness in diverse scenarios. The average value comparison of different competing techniques is also shown in Fig. 5.

VI. DISCUSSION

The comparative evaluation conducted in this study aimed to assess the effectiveness of different approaches—COA-ANN, GWO-ANN, PSO-ANN, and BP-ANN—across three distinct datasets. The assessment was based on a range of evaluation metrics, including accuracy, F1-Score, precision, and recall. The results presented in Tables I, II, and III reveal valuable insights into the strengths and limitations of each approach, shedding light on their performance across varying datasets. Notably, the variation in performance across datasets underscores the influence of data complexity on technique selection. COA-ANN consistently exhibited exceptional performance across all datasets, indicating its ability to capture intricate patterns effectively. GWO-ANN displayed consistent and competitive performance, suggesting its potential for generalization in diverse scenarios but slightly lagged behind the COA-ANN. In contrast, PSO-ANN showcased sensitivity to dataset characteristics, as it demonstrated competitive performance in some instances while faltering in others. Meanwhile, BP-ANN consistently lagged in terms of accuracy and F1-Score, emphasizing the limitations of traditional gradient-based optimization methods in addressing intricate relationships within data. These results collectively highlight the critical role that both optimization algorithms and neural network architectures play in determining the success of a given approach.

CONCLUSION

In this paper, a novel bio-inspired optimization algorithm is used within a neural network architecture. The objective of this study was to improve the predictive accuracy of Cardiovascular Diseases by incorporating the coati optimization algorithm in the artificial neural network. A 3-layered neural network was used in this study with the hidden layer having 10 neurons. The sigmoid activation function and quadratic cost function were employed in our framework. By incorporating two datasets related to the heart and one related to stroke disease, our objective was to access the generalization ability of COA-ANN. The results obtained by using COA-ANN were commendable as the predictive accuracy average outperformed the other three networks namely; Particle Swarm Optimizer based ANN (PSO-ANN), Grey Wolf Optimizer based ANN

(GWO-ANN), backpropagation based ANN (BP-ANN). The COA-ANN had the highest predictive accuracy average of 98.43%, surpassing GWO-ANN, PSO-ANN, and BP-ANN. The second best average was obtained by GWO-ANN with an average of 95.79%. The results have shown that the COA-ANN outperforms the other networks by more than 2% by utilizing the same computational resources. A deep neural network (DNN) architecture with transfer learning techniques can be used in the future to further enhance the generalization ability and performance of AI models for more complex CVD datasets.

Acknowledgment

This research is supported by Top Research Center Mechatronics (TRCM), UiA, Norway.

References

[1] World Health Organization. “Cardiovascular diseases (cvds).” Accessed on 10 Aug 2023. (), [Online]. Available: https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab_1 (visited on 03/10/2023).

[2] A. K. Malakar, D. Choudhury, B. Halder, P. Paul, A. Uddin, and S. Chakraborty, “A review on coronary artery disease, its risk factors, and therapeutics,” *Journal of cellular physiology*, vol. 234, no. 10, pp. 16 812–16 823, 2019.

[3] World Health Organization. “The top 10 causes of death.” Accessed on 10 Aug 2023. (Dec. 9, 2020), [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death (visited on 07/10/2023).

[4] P. Rothwell, A. Coull, L. Silver, *et al.*, “Population-based study of event-rate, incidence, case fatality, and mortality for all acute vascular events in all arterial territories (oxford vascular study),” *The Lancet*, vol. 366, no. 9499, pp. 1773–1783, 2005.

[5] K. W. Johnson, J. Torres Soto, B. S. Glicksberg, *et al.*, “Artificial intelligence in cardiology,” *Journal of the American College of Cardiology*, vol. 71, no. 23, pp. 2668–2679, 2018.

[6] I. Guyon and A. Elisseeff, “An introduction to variable and feature selection,” *Journal of machine learning research*, vol. 3, no. Mar, pp. 1157–1182, 2003.

[7] T. O. Ayodele, “Types of machine learning algorithms,” *New advances in machine learning*, vol. 3, pp. 19–48, 2010.

[8] S. K. R. Moosavi, A. Saadat, Z. Abaid, W. Ni, K. Li, and M. Guizani, “Feature selection based on dataset variance optimization using hybrid sine cosine – firehawk algorithm (hscfha),” *SSRN*, 2023, Available at SSRN: http://dx.doi.org/10.2139/ssrn.4485391, Submitted to FGCS.

[9] A. Exposito-Marquez, C. Exposito-Izquierdo, J. Brito-Santana, and J. A. Moreno-Pérez, “Greedy randomized adaptive search procedure to design waste collection routes in la palma,” *Computers & Industrial Engineering*, vol. 137, p. 106 047, 2019.

[10] H.-G. Beyer and H.-P. Schwefel, “Evolution strategies–a comprehensive introduction,” *Natural computing*, vol. 1, pp. 3–52, 2002.

[11] I. Sharma, V. Kumar, and S. Sharma, “A comprehensive survey on grey wolf optimization,” *Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science)*, vol. 15, no. 3, pp. 323–333, 2022.

[12] A. G. Gad, “Particle swarm optimization algorithm and its applications: A systematic review,” *Archives of computational methods in engineering*, vol. 29, no. 5, pp. 2531–2561, 2022.

[13] M. Morin, I. Abi-Zeid, and C.-G. Quimper, “Ant colony optimization for path planning in search and rescue operations,” *European Journal of Operational Research*, vol. 305, no. 1, pp. 53–63, 2023.

[14] A. M. Fathollahi-Fard, M. Hajiaghaei-Keshteli, and R. Tavakkoli-Moghaddam, “Red deer algorithm (rda): A new nature-inspired metaheuristic,” *Soft computing*, vol. 24, pp. 14 637–14 665, 2020.

[15] M. Dehghani, Z. Montazeri, E. Trojovská, and P. Trojovskỳ, “Coati optimization algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems,” *Knowledge-Based Systems*, vol. 259, p. 110 011, 2023.

[16] E. A. A. Shenouda, “A quantitative comparison of different mlp activation functions in classification,” in *International Symposium on Neural Networks*, Springer, 2006, pp. 849–857.

[17] J.-W. Lin and J.-S. Chiou, “Active probability backpropagation neural network model for monthly prediction of probabilistic seismic hazard analysis in taiwan,” *IEEE Access*, vol. 7, pp. 108 990–109 014, 2019.

[18] V. M. Joy and S. Krishnakumar, “Optimal design of adaptive power scheduling using modified ant colony optimization algorithm.,” *International Journal of Electrical & Computer Engineering (2088-8708)*, vol. 10, no. 1, 2020.