

Received February 2, 2021, accepted February 16, 2021, date of publication February 22, 2021, date of current version March 8, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3061058

Advanced Meta-Heuristics, Convolutional Neural Networks, and Feature Selectors for Efficient COVID-19 X-Ray Chest Image Classification

EL-SAYED M. EL-KENAWY¹, (Member, IEEE),
SEYEDALI MIRJALILI^{2,3}, (Senior Member, IEEE),
ABDELHAMEED IBRAHIM⁴, (Member, IEEE), MOHAMMED ALRAHMAWY⁵,
M. EL-SAID^{6,7}, ROKAIA M. ZAKI^{1,8}, AND MARWA METWALLY EID¹

¹Department of Communications and Electronics, Delta Higher Institute of Engineering and Technology (DHJET), Mansoura 35111, Egypt

²Centre for Artificial Intelligence Research and Optimization, Torrens University Australia, Fortitude Valley, QLD 4006, Australia

³Yonsei Frontier Lab, Yonsei University, Seoul 03722, South Korea

⁴Computer Engineering and Control Systems Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt

⁵Department of Computer Science, Faculty of Computers and Information, Mansoura University, Mansoura 35516, Egypt

⁶Electrical Engineering Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt

⁷Delta Higher Institute of Engineering and Technology (DHJET), Mansoura 35111, Egypt

⁸Department of Electrical Engineering, Shoubra Faculty of Engineering, Benha University, Benha 11629, Egypt

Corresponding authors: Abdelhameed Ibrahim (afai79@mans.edu.eg) and El-Sayed M. El-Kenawy (skenawy@ieee.org)

ABSTRACT The chest X-ray is considered a significant clinical utility for basic examination and diagnosis. The human lung area can be affected by various infections, such as bacteria and viruses, leading to pneumonia. Efficient and reliable classification method facilitates the diagnosis of such infections. Deep transfer learning has been introduced for pneumonia detection from chest X-rays in different models. However, there is still a need for further improvements in the feature extraction and advanced classification stages. This paper proposes a classification method with two stages to classify different cases from the chest X-ray images based on a proposed Advanced Squirrel Search Optimization Algorithm (ASSOA). The first stage is the feature learning and extraction processes based on a Convolutional Neural Network (CNN) model named ResNet-50 with image augmentation and dropout processes. The ASSOA algorithm is then applied to the extracted features for the feature selection process. Finally, the Multi-layer Perceptron (MLP) Neural Network's connection weights are optimized by the proposed ASSOA algorithm (using the selected features) to classify input cases. A Kaggle chest X-ray images (Pneumonia) dataset consists of 5,863 X-rays is employed in the experiments. The proposed ASSOA algorithm is compared with the basic Squirrel Search (SS) optimization algorithm, Grey Wolf Optimizer (GWO), and Genetic Algorithm (GA) for feature selection to validate its efficiency. The proposed (ASSOA + MLP) is also compared with other classifiers, based on (SS + MLP), (GWO + MLP), and (GA + MLP), in performance metrics. The proposed (ASSOA + MLP) algorithm achieved a classification mean accuracy of (99.26%). The ASSOA + MLP algorithm also achieved a classification mean accuracy of (99.7%) for a chest X-ray COVID-19 dataset tested from GitHub. The results and statistical tests demonstrate the high effectiveness of the proposed method in determining the infected cases.

INDEX TERMS Chest X-ray, transfer learning, convolutional neural network, squirrel search optimization, multilayer perceptron, optimization algorithm.

I. INTRODUCTION

Medical images are teeming with many features that can be considered for inspection. Generally, many processes in

The associate editor coordinating the review of this manuscript and approving it for publication was Krishna Kant Singh¹.

Computer-Aided System (CAD), such as pre-processing, isolating Regions of Interest (ROIs), and feature extracting process, can help to get the accurate classification of the diseases [1]. There are various approaches for highlighting ROIs, extracting the salient features, and suppressing the associated noises [2]–[4]. Rule-based techniques have

limited performance, and to improve efficiency, they are usually consolidated. Traditional approaches focused on fetching geometric or handcrafted features are generally treated to reduce dimensionality, elapsed time, and redundancy features concerning extract salient features. Moreover, these methods suffer from failures that affect the classification accuracy. Hence, any improvements in the feature extraction step and the segmentation process are required [5]–[8]. For image classification tasks, some traditional classification methods have also achieved excellent results in recent years [9]–[12]; however, the deep learning methods have some advantages over the traditional methods. CNN's (Convolutional Neural Networks) or pre-trained networks are commonly involved in different medical imaging tasks. They can offer rather good performance in analyzing high-resolution images as in abdominal X-rays. However, a need for sufficient amounts of the training dataset is a critical problem [13]–[15].

More training phases with large datasets are mainly required to apply the pre-trained networks for medical image classification tasks. Hence, in some cases, adopting these types of networks as classifiers are often not the preferred method to apply the CNN to the diagnosis tasks of CAD [16]–[18]. Using pre-trained networks (e.g., DCNNs), extra complicated datasets such as pneumonia's presence/absence did not seem good. Therefore, more data augmentation samples and training can improve the efficiency [19], [20]. While identification systems based on CNN provide greater precision for various tasks, the key downside of these methods is the need for heavy training [21]–[23]. Hence, machine learning increases rapidly, which has caused many technical breakthroughs and is extensively employed in many fields. As a significant part of machine learning tasks, optimization has attracted much attention in several research areas. With the expedited growth of the amount of employed data and the increase of design complexity, optimization approaches in machine learning face more further challenges.

For a specific problem, optimization can be the most reliable solution between all available solutions, especially towards multi-dimensional space [24]. Practically, this involves maximization or minimization of an objective function. The objective function defines the solution candidate's quality and efficiency represented by a particular vector in a search domain. There are two classes of optimization: nonlinear versus linear [25], [26]. Meta-heuristic algorithms are considered among the most powerful methods for solving real-world engineering problems [27]. Most of these algorithms' derivation is done from physical algorithms' rational behavior in nature, biological inspired algorithms' behavior, swarm particles' collective intelligence, and evolutionary algorithms' fittest theory of survival [28], [29]. These optimization techniques provide acceptable solutions in a reasonable time with less computational effort. They are mostly used in engineering and science for finding solutions to complex and challenging problems because: a) of their utilization in different issues that come under other subjects, b) of no requirement of gradient information,

c) they can bypass local optima, and d) they are easy to be implemented and are dependent on comparatively simple concepts.

This paper proposes a classification method to classify infected cases from the chest X-ray images. The method can decrease detection costs significantly. First, a feature learning stage is developed using the CNN model named ResNet-50 with image augmentation as a pre-processing and dropout as post-processing. Second, the features are extracted to start the feature selection process. A proposed Advanced Squirrel Search Optimization Algorithm (ASSOA) is developed for feature selection. The advanced classification stage starts to classify the infected cases using the optimized Multilayer Perceptron Neural Network (MLP) by the proposed ASSOA algorithm. ASSOA's basic rule in the classification stage is to optimize the connection weights of MLP to improve accuracy. A dataset from Kaggle, chest X-ray images (Pneumonia) dataset [30] consists of 5,863 X-ray images used in experiments. A chest X-ray COVID-19 dataset [31] is also tested in the experiments. The proposed ASSOA algorithm is compared with the basic Squirrel Search (SS) optimization algorithm [32], Grey Wolf Optimizer (GWO) [29], and Genetic Algorithm (GA) [33] for feature selection to test its efficiency. The ASSOA + MLP algorithm is also compared with other classifiers, based on (SS + MLP), (GWO + MLP), and (GA + MLP), in performance metrics. Moreover, Wilcoxon rank-sum and one-way analysis of variance (ANOVA) are tested to statistically verify the proposed algorithm's superiority.

This paper's main contributions are as follows:

- An Advanced Squirrel Search Optimization Algorithm (ASSOA) is developed for feature extraction and classification.
- The proposed ASSOA algorithm adds horizontal, vertical, diagonal, and exponential movements to the basic moves in the search process of the basic SS algorithm.
- A new agents' relocation equation is modeled in the proposed ASSOA algorithm, affecting local and global optima under specific conditions.
- A classification method for chest X-ray images is proposed based on the ASSOA algorithm.
- The classification method is tested using a dataset from Kaggle with 5,863 chest X-ray images.
- The classification method is also tested for a chest X-ray COVID-19 dataset from GitHub.
- Wilcoxon rank-sum and ANOVA statistical tests are performed to ensure the proposed ASSOA algorithm quality.

The next sections of this paper are as follows: Section II presents the related works. The materials and methods used in the study are defined in Section III. In-depth, Section IV describes the proposed method and the ASSOA algorithm. The experimental results are shown in Section V. Section IV discusses the proposed method findings. The research conclusions are seen in section VII.

TABLE 1. Recent machine learning research for classification of CT and X-ray images.

Reference	Methods	# of samples	# of classes	Type of Images	Accuracy	AUC
X. Wang <i>et al.</i> (2020) [52]	Weakly supervised deep learning model	450	3	CT	96.2%	0.970
K. Zhang <i>et al.</i> (2020) [48]	AI system based on ResNet-18	3,777	3	CT	92.49%	0.981
C. Butt <i>et al.</i> (2020) [49]	Multiple CNN models based on ResNet-18	618	3	CT	86.7%	0.996
O. Stephen <i>et al.</i> (2019) [34]	Efficient Deep Learning Approach	200	2	X-ray	93.73%	—
D. Varshni <i>et al.</i> (2019) [46]	DenseNet-169 + SVM	2862	2	X-ray	—	0.8
H. Panwar <i>et al.</i> (2020) [50]	nCOVnet, transfer learning, deep CNN	337	2	X-ray	88.10%	0.881
M. Nour <i>et al.</i> (2020) [51]	Training CNN model, feature extraction, SVM	2,905	3	X-ray	98.97%	0.994
P. Chhikara <i>et al.</i> (2020) [47]	Deep CNN with Transfer Learning	5866	2	X-ray	90.16%	0.911
M. A. Elaziz <i>et al.</i> (2020) [53]	MRFODE feature selection method, KNN classifier	1891	2	X-ray	96.09%	—
		1560	2	X-ray	98.09%	—

II. RELATED WORK

Large datasets availability and the recent advances in deep learning models have led to the possession of power-assisted algorithms, which beats the medical professionals in various clinical image resolution. These images are such as cancer classification [34], detection of arrhythmia [35], [36], identification of haemorrhage [37], and diagnosis/detection of diabetic retinopathy [38]. Using radiography, the automated diagnosis of chest diseases has gained a lot of enthusiasm and interest. Several CNN models' efficiency on various oddities certainly does not do well with all abnormalities, deep-learning approaches, and ensemble models may improve classification accuracy considerably reviewed to other Versions. Statistical dependence was studied between the precision levels of the predictions and the Multi-label Disease Classification (MDC). In the literature, the detection of health conditions from chest X-ray images was performed based on different methods [39]–[41]. The processes for X-ray cardiovascular angiogram images are proposed in the literature [42], [43].

Recent research has implemented several automatic pneumonia detection systems based on chest X-rays [44], [45]. Deep learning is applied for the training AI algorithms to detect pneumonia by studying chest X-ray images [46]. In terms of accuracy, Chhikara *et al.* achieved an accuracy of (90.1%) in [47] using 5,866 chest X-ray images compared to the latest models of classification. The CNN model proposed by Okeke Stephen *et al.* in [34], was constructed by extracting characteristics from the images of chest X-ray to test the existence of pneumonia. The authors in that model deployed multiple data augmentation algorithms to enhance both validation and classification accuracy of their model to achieve an accuracy of (93.73%).

An AI approach to diagnosing COVID-19 and other types of pneumonia is already developed in [48]. For COVID-19, their proposed method achieved an AUC (area under the curve) of (0.981) and accuracy of (92.49%). Butt *et al.* in [49] A CNN model called ResNet-18 was proposed to classify the CT images as COVID-19, regular, and pneumonia. With an AUC value of (0.996), they can achieve an accuracy of (86.7%). Authors in [50] The nCOVnet, based on deep learning, was proposed to detect COVID-19 by analyzing patients' X-ray images.

Their nCOVnet system obtained an AUC of (0.881) and a COVID-19 accuracy of (88.10%).

Nour *et al.* [51] using X-ray images, a CNN model trained from scratch was suggested. The model's extracted features fed K-NN, SVM, and decision tree in their model. The SVM classifier achieved an accuracy of (98.97%). Hu *et al.* in [52] A weakly-supervised CNN model was proposed, which achieved an accuracy of (96.2%) with an AUC value of (0.970). To classify the chest's x-ray images into COVID-19 or non-COVID-19, an ML-method in [53] was proposed. A Manta-Ray Foraging Optimization technique, using differential evolution, was developed for feature selection. The authors evaluated their method by testing two COVID-19 x-ray datasets. The recent machine learning research for CT and X-ray images is summarized in Table 1.

Therefore, building a classification method for various infections is one of the most critical issues prohibitively expensive for mass adoption. Deep transfer learning has been introduced for pneumonia detection from chest X-rays in different literature models. However, there is still a need for more improvements in the feature extraction and classification stages.

III. MATERIALS AND METHODS

This section introduces the chest X-ray datasets used in this paper and will also discuss the essential CNN deep transfer learning, multilayer perceptron neural network, and the original Squirrel Search (SS) optimization algorithm.

A. DATASETS

Chest (Pneumonia) X-ray images from Kaggle dataset [30] has been used. In (JPEG) format, the dataset has 5,863 X-rays. It is classified into two cases, either normal or pneumonia. In this paper, The Kaggle dataset has been selected because it is used in many forms of research globally and makes comparisons that can enrich scientific research. Figure 1(a) shows image samples of normal Pneumonia-free cases, while Fig. 1(b) and Fig. 1(c) present Pneumonia image samples (Bacteria and Viral cases), respectively. Another chest X-ray COVID-19 dataset [31] is also tested in the experiments and image samples are shown in Fig. 2. Besides indirect collection from hospitals and physicians, the COVID-19 dataset is obtained from public

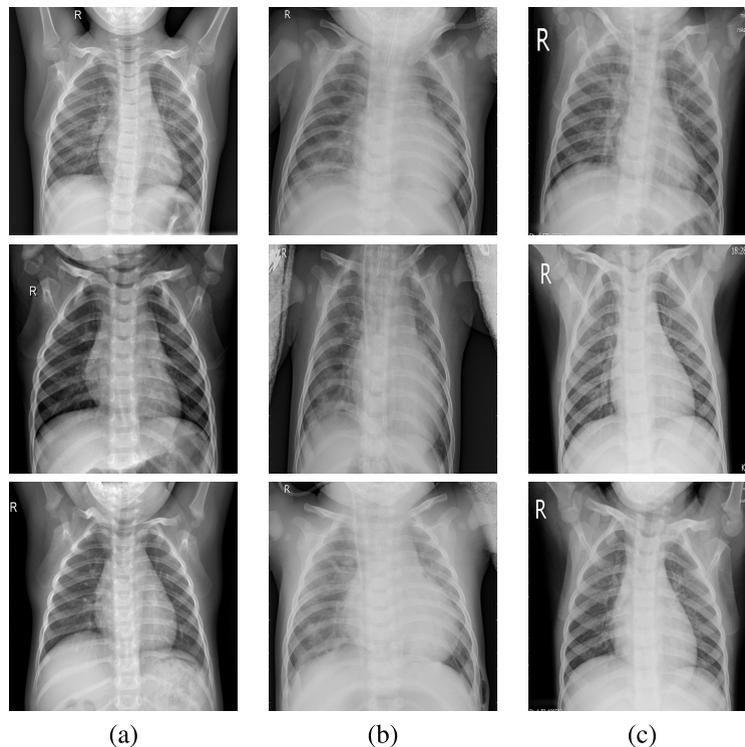


FIGURE 1. Samples of the original chest X-ray images [30]; (a) Normal cases, (b) Bacteria cases, and (c) Viral cases.

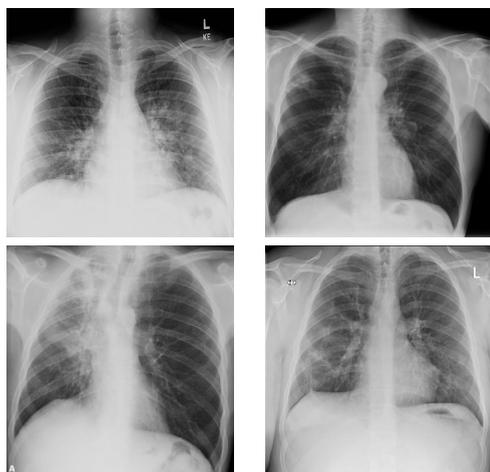


FIGURE 2. Samples of the original chest X-ray COVID-19 infected cases [31] tested in the experiments.

sources. In the GitHub repo, all data and images are released publicly. The tested dataset's project was accepted by the Ethics Committee of the University of Montreal #CERSES-20-058-D.

B. DEEP TRANSFER LEARNING AND CNN

In traditional learning, the model is isolated and based mainly on specific tasks and particular datasets [54]. The knowledge, in this learning, cannot be transferred from one model to

another. In the transfer learning, knowledge, such as features and weights, can be transferred from the pre-trained model to new training models and different problems that may have fewer data. Transfer Learning is usually applied in various models for a dataset with less data than the dataset used to train the model. Multitask learning allows several tasks to learn simultaneously, which can help the model receive multiple tasks at once. The learner initially may have no idea about the target task [47].

CNN's [55], a deep neural network, is known to be ideal for image processing applications and can achieve greater precision in the subject of disease classification than conventional approaches. It can thus be used in applications such as clustering, detecting objects, and classifying images. Several CNN models have recently been introduced, such as AlexNet, [56], VGGNet [57], GoogLeNet [58], Spotmole [59] and ResNet [60]. Convolution models used in the CNN models have different layers; higher classification accuracy is achieved if the number of convolution layers increases. [61]–[63].

Residual Network (ResNet) is known as an efficient CNN model [60]. The ResNet model was declared in 2016 to be the best paper at the Computer Vision and Pattern Recognition Conference (CVPR 06). The ResNet concept is based on the assumption that only a residual correction of the previous layer should be a deeper network training that can function efficiently, not transforming the whole feature space. The main idea of ResNets is not to learn

the mapping from $x \rightarrow F(x)$, but instead learns the mapping from $x \rightarrow F(x) + G(x)$. Thus, if output $F(x)$, for input x , have the same dimension, $G(x) = x$ function is identity and the connection is identity. ResNet, without exploding and gradient vanishing issues, has much deeper neural network training.

C. MULTILAYER PERCEPTRON

Feed-forward neural networks are considered supervised machine learning methods consisting of neurons distributed over fully connected layers. The first (input) layer maps the network input variables, and the last layer is the output one. Layers between the first and last layers are called hidden layers [64], [65]. Multilayer perceptron (MLP) is a common type of feed-forward network. The neurons interconnect in MLP, are one-directional fashion. The weights of the connections are within $[-1, 1]$. Figure 3 shows the MLP network, which includes one layer between input and output. To calculate the node output value, the weighted sum is firstly calculated as follows:

$$S_j = \sum_{i=1}^n w_{ij}I_i + \beta_j \tag{1}$$

where I_i represents input variable i and w_{ij} indicates connection weight between I_i and neuron j in the hidden layer. β_j is bias value for this layer. By applying the sigmoid activation function, which is the mostly applied, node j output is defined as

$$f_j(S_j) = \frac{1}{1 + \exp^{-S_j}} \tag{2}$$

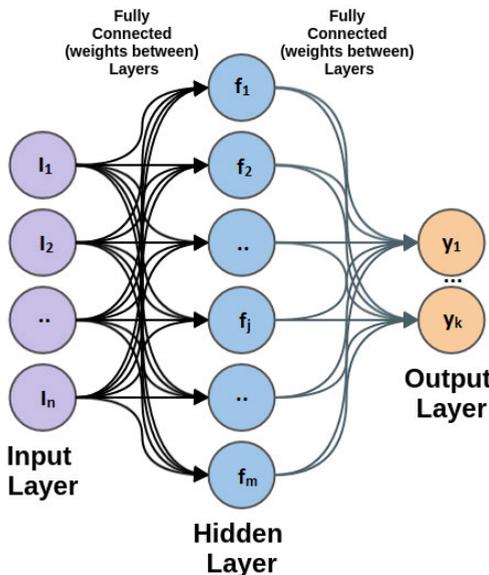


FIGURE 3. Neural Network - Multilayer Perceptron (MLP).

Based on the value of $f_j(S_j)$ for all hidden layer neurons, the following equation can define the network output:

$$y_k = \sum_{j=1}^m w_{jk}f_j(S_j) + \beta_k \tag{3}$$

where w_{jk} indicates weights between neuron j in the hidden layer and output node k and β_k is the bias value for the output layer.

D. SQUIRREL SEARCH BASIC OPTIMIZATION ALGORITHM

The Squirrel Search (SS) basic optimization algorithm simulates the search process of flying squirrels [32]. The SS algorithm considers that the squirrels are moving between three kinds of trees named normal, oak, and hickory trees. The oak and hickory trees are the nuts food source, while normal trees have no food source. Mathematically, the SS algorithm assumes the squirrels are flying to search for three oak trees and one hickory tree as nutritious food resources N_{fs} available for n flying squirrels (FS).

The flying agents' locations is in matrix form as follows:

$$FS = \begin{bmatrix} FS_{1,1} & FS_{1,2} & FS_{1,3} & \dots & FS_{1,d} \\ FS_{2,1} & FS_{2,2} & FS_{2,3} & \dots & FS_{2,d} \\ FS_{3,1} & FS_{3,2} & FS_{3,3} & \dots & FS_{3,d} \\ \dots & \dots & \dots & \dots & \dots \\ FS_{n,1} & FS_{n,2} & FS_{n,3} & \dots & FS_{n,d} \end{bmatrix} \tag{4}$$

where $FS_{i,j}$ indicates i^{th} flying squirrel in the j^{th} dimension for $i \in 1, 2, 3, \dots, n$ and $j \in 1, 2, 3, \dots, d$. The initial locations of $FS_{i,j}$ are uniform distribution within lower and upper bounds. The fitness values $f = f_1, f_2, f_3, \dots, f_n$ are calculated for each flying squirrel as in the following array

$$f = \begin{bmatrix} f_1(FS_{1,1}, FS_{1,2}, FS_{1,3}, \dots, FS_{1,d}) \\ f_2(FS_{2,1}, FS_{2,2}, FS_{2,3}, \dots, FS_{2,d}) \\ f_3(FS_{3,1}, FS_{3,2}, FS_{3,3}, \dots, FS_{3,d}) \\ \dots \\ f_n(FS_{n,1}, FS_{n,2}, FS_{n,3}, \dots, FS_{n,d}) \end{bmatrix} \tag{5}$$

where the fitness value indicates the food source quality searched by each flying squirrel. The optimal value means a hickory tree. These values are then reordered in ascending order. The first best solution is declared to be FS_{ht} on the hickory nut tree followed by three best solutions that are considered to be FS_{at} on the acorn nuts trees. The remaining solutions are supposed to be FS_{nt} on normal trees.

New location generation mathematically for each flying squirrel is considered as one of the three following cases:

Case 1: Location of FS_{at} and moving to the hickory nut tree:

$$FS_{at}^{t+1} = \begin{cases} FS_{at}^t + d_g \times G_c(FS_{ht}^t - FS_{at}^t) & \text{if } R_1 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \tag{6}$$

Case 2: Location of FS_{nt} and moving to the acorn nut trees:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + d_g \times G_c(FS_{at}^t - FS_{nt}^t) & \text{if } R_2 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \tag{7}$$

Case 3: Location of FS_{nt} and moving to the hickory nut tree:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + d_g \times G_c(FS_{ht}^t - FS_{nt}^t) & \text{if } R_3 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \quad (8)$$

where R_1, R_2 , and R_3 are random numbers $\in [0, 1]$. The d_g parameter is random distance for gliding and t indicates the current iteration. G_c is equal to 1.9 and it is constant to achieve exploration and exploitation balance, and the value of P_{dp} probability is equal to 0.1 for the three cases.

The seasonal constant (S_c) is calculated from the following equation to check the monitoring condition ($S_c^t < S_{min}$) as

$$S_c^t = \sqrt{\sum_{k=1}^d (FS_{at,k}^t - FS_{ht,k}^t)^2}, \quad t = 1, 2, 3. \quad (9)$$

$$S_{min} = \frac{10E^{-6}}{(365)^{t/(t_m/2.5)}} \quad (10)$$

where t is the current iteration and t_m represents the maximum iteration value. The value of S_{min} can affect the algorithm exploration and exploitation capabilities during iterations. If specific condition is occurred, such flying squirrels's relocation is modeled by Eq. 11

$$FS_{nt}^{new} = FS_L + Levy(n) \times (FS_U - FS_L) \quad (11)$$

where the distribution $Levy$ helps in encouraging better search space exploration. The calculation of the $Levy$ flight is as follows:

$$Levy(x) = 0.01 \times \frac{r_a \times \sigma}{|r_b|^{\frac{1}{\beta}}} \quad (12)$$

where the parameters r_a and r_b are random in $[0, 1]$. β is equal to 1.5 in the SS algorithm and σ is calculated as

$$\sigma = \left(\frac{\Gamma(1 + \beta) \times \sin(\frac{\pi\beta}{2})}{\Gamma(\frac{1+\beta}{2}) \times \beta \times 2^{(\frac{\beta-1}{2})}} \right)^{1/\beta} \quad (13)$$

where $\Gamma(x) = (x - 1)!$. The basic Squirrel Search (SS) optimization algorithm is explained step by step in Algorithm 1.

IV. PROPOSED CLASSIFICATION METHOD

The proposed classification method consists of two stages. The first stage has a feature engineering process, including image augmentation, CNN training using the ResNet-50 model, transfer learning, and dropout. The proposed ASSOA algorithm is then applied to select features from the ResNet-50 model's extracted features. The second stage involves the classification process to classify cases in which the MLP is optimized by the proposed ASSOA algorithm (ASSOA + MLP).

A. FEATURE ENGINEERING STAGE

The ResNet-50 model is applied in this stage as a part of the proposed method for features extraction from the chest X-ray images in the fully connected layer by altering the nodes

Algorithm 1 Basic SS Optimization Algorithm [32]

- 1: **Initialize** SS population $FS_i (i = 1, 2, \dots, n)$ with size n using Eq. (4), maximum iterations t_m , and fitness function F_n .
- 2: **Initialize** SS parameters $R_1, R_2, R_3, n_1, n_2, n_3, P_{dp}, G_c, t = 1$
- 3: **Calculate** fitness function F_n for each FS_i using Eq. (5)
- 4: **Sort** flying squirrels locations in ascending order
- 5: **Find** the first best individual FS_{ht}
- 6: **Find** the next three best individuals FS_{at}
- 7: **Find** the normal individuals FS_{nt}
- 8: **while** $t \leq t_m$ (Stopping condition) **do**
- 9: **for** $(t = 1 : n_1)$ **do**
- 10: **if** $(R_1 \geq P_{dp})$ **then**
- 11: $FS_{at}^{t+1} = FS_{at}^t + d_g \times G_c(FS_{ht}^t - FS_{at}^t)$
- 12: **else**
- 13: $FS_{at}^{t+1} = \text{Random location}$
- 14: **end if**
- 15: **end for**
- 16: **for** $(t = 1 : n_2)$ **do**
- 17: **if** $(R_2 \geq P_{dp})$ **then**
- 18: $FS_{nt}^{t+1} = FS_{nt}^t + d_g \times G_c(FS_{at}^t - FS_{nt}^t)$
- 19: **else**
- 20: $FS_{nt}^{t+1} = \text{Random location}$
- 21: **end if**
- 22: **end for**
- 23: **for** $(t = 1 : n_3)$ **do**
- 24: **if** $(R_3 \geq P_{dp})$ **then**
- 25: $FS_{nt}^{t+1} = FS_{nt}^t + d_g \times G_c(FS_{ht}^t - FS_{nt}^t)$
- 26: **else**
- 27: $FS_{nt}^{t+1} = \text{Random location}$
- 28: **end if**
- 29: **end for**
- 30: **Calculate** seasonal constant (S_c^t) using Eq. (9)
- 31: **Calculate** minimum value of seasonal constant (S_{min}) using Eq. (10)
- 32: **if** $(S_c^t < S_{min})$ **then**
- 33: $FS_{nt}^{new} = FS_L + Levy(n) \times (FS_U - FS_L)$
- 34: **end if**
- 35: **Update** S_{min} using Eq. (10)
- 36: **Set** $t = t + 1$
- 37: **end while**
- 38: **Return** optimal solution FS_{ht}

and doing a fine-tuning based on the input dataset. Each input image is resized to 224×224 pixels to be suitable for the model. Then, the Min-Max-Scalar is used to normalize the i th input image x_i to a scale from 0 to 1 by applying the following equation.

$$x'_i = \frac{x_i - \min(x_i)}{\max(x_i) - \min(x_i)} \quad (14)$$

After the resizing and normalization, the output image x'_i is used as input to the CNN model. The adopted CNN structure of the number of filters and layers and the related specifi-

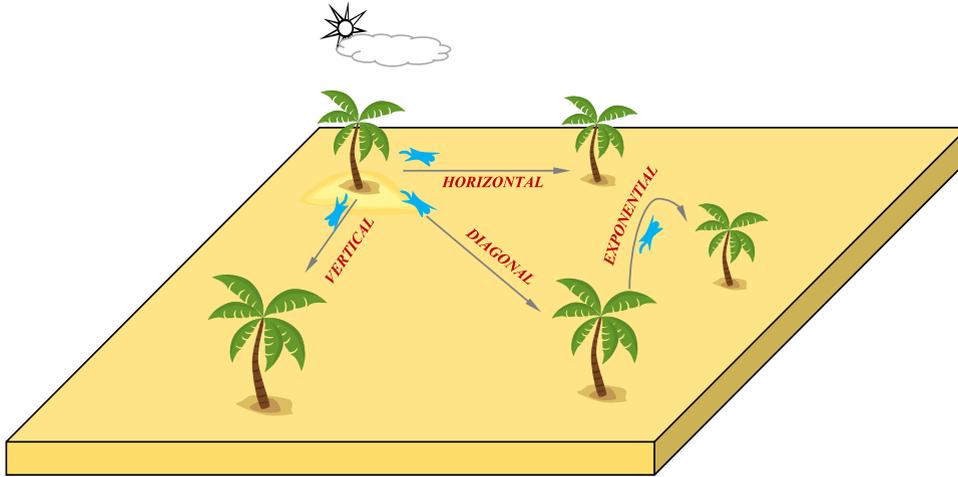


FIGURE 4. Squirrel movements in the Proposed Advanced Squirrel Search Optimization Algorithm (ASSOA).

cations are identical to the ResNet-50 model. This model focuses on classifying input case categories. To reduce the overfitting problem during network learning, two regularization techniques of dropout and image augmentation have been applied in this research. The dropout is applied during the training procedure of CNN, and image augmentation [66] is used for the X-ray images' input images. Data Augmentation is applied to improve the quality and size of the training datasets.

B. THE PROPOSED ASSOA ALGORITHM

The proposed ASSOA algorithm adds horizontal, vertical, diagonal, and exponential movements to the basic moves in the search process of flying squirrels, as shown in Fig. 4. The ASSOA algorithm considers, as in the basic SS algorithm, that the squirrels are moving between three kinds of trees named normal, oak, and hickory trees. The nuts food sources are the oak and hickory trees, while there are no food sources on the other trees.

Mathematically, the ASSOA algorithm assumes the squirrels are flying in directions shown in Fig. 4 to search for one hickory tree, the best solution, and three oak trees, next best solutions, as nutritious food resources N_{fs} available for n flying squirrels (FS). The following matrices represent the flying squirrels' locations and velocities:

$$FS = \begin{bmatrix} FS_{1,1} & FS_{1,2} & FS_{1,3} & \dots & FS_{1,d} \\ FS_{2,1} & FS_{2,2} & FS_{2,3} & \dots & FS_{2,d} \\ FS_{3,1} & FS_{3,2} & FS_{3,3} & \dots & FS_{3,d} \\ \dots & \dots & \dots & \dots & \dots \\ FS_{n,1} & FS_{n,2} & FS_{n,3} & \dots & FS_{n,d} \end{bmatrix} \quad (15)$$

$$V = \begin{bmatrix} V_{1,1} & V_{1,2} & V_{1,3} & \dots & V_{1,d} \\ V_{2,1} & V_{2,2} & V_{2,3} & \dots & V_{2,d} \\ V_{3,1} & V_{3,2} & V_{3,3} & \dots & V_{3,d} \\ \dots & \dots & \dots & \dots & \dots \\ V_{n,1} & V_{n,2} & V_{n,3} & \dots & V_{n,d} \end{bmatrix} \quad (16)$$

where $FS_{i,j}$ indicates i^{th} flying squirrel location in the j^{th} dimension for $i \in 1, 2, 3, \dots, n$ and $j \in 1, 2, 3, \dots, d$. $V_{i,j}$ indicates i^{th} flying squirrel velocity in the j^{th} dimension for $i \in 1, 2, 3, \dots, n$ and $j \in 1, 2, 3, \dots, d$. The initial locations of $FS_{i,j}$ are uniform distribution within lower and upper bounds. The fitness values $f = f_1, f_2, f_3, \dots, f_n$ are calculated for each flying squirrel as in Eq. 5. The optimal value means a hickory tree. These values are then sorted in ascending order. The first best solution is declared to be FS_{ht} on the hickory nut tree followed by three best solutions that are considered to be FS_{at} on the acorn nuts trees. The remaining solutions are supposed to be FS_{nt} on normal trees.

In the ASSOA algorithm, the new location generation for each flying squirrel is updated as in the following cases. For a random value p , the following cases will be applied if $p \geq 0.5$:

Case 1: Location of FS_{at} and moving to the hickory nut tree:

$$FS_{at}^{t+1} = \begin{cases} FS_{at}^t + d_g \times G_c (FS_{ht}^t - FS_{at}^t) & \text{if } R_1 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \quad (17)$$

Case 2: Location of FS_{nt} and moving to the acorn nut trees:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + d_g \times G_c (FS_{at}^t - FS_{nt}^t) & \text{if } R_2 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \quad (18)$$

Case 3: Location of FS_{nt} and moving to the hickory nut tree:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + d_g \times G_c (FS_{ht}^t - FS_{nt}^t) & \text{if } R_3 \geq P_{dp} \\ \text{Random location} & \text{otherwise} \end{cases} \quad (19)$$

where R_1 , R_2 , and R_3 are random numbers $\in [0, 1]$. The d_g parameter is random distance for gliding and t indicates the current iteration. G_c is equal to 1.9 and it is constant to achieve the exploration and exploitation balance, and the value of P_{dp} probability is equal to 0.1 for the three cases.

For the random value p , the following cases will be applied if $p < 0.5$:

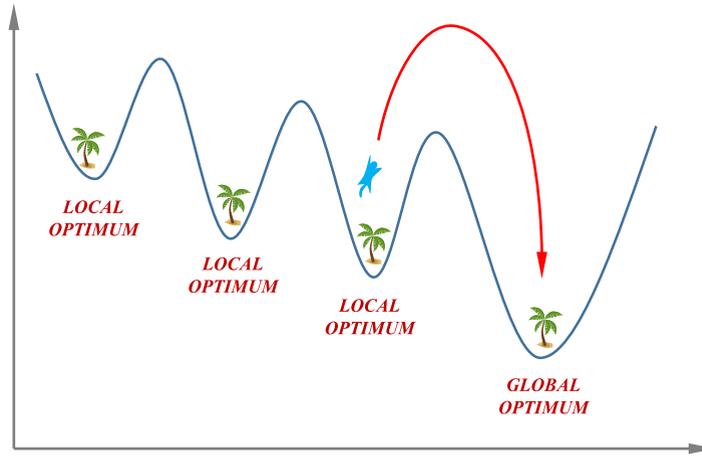


FIGURE 5. Local and global optima of the proposed ASSOA algorithm.

Case 4: Location of FS_{nt} and moving diagonally:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + V_{nt}^t + \\ c_1 r(FS_{ht}^t - FS_{nt}^t) + \\ c_2 r(FS_{at}^t - FS_{nt}^t) & \text{if } P_a < a \\ \text{Random } FS_{rand}^t \in FS_{nt}^t & \text{otherwise} \end{cases} \quad (20)$$

where c_1, c_2, r, P_a , and a are random numbers $\in [0, 1]$. In case of choosing a random agent FS_{rand}^t from the normal agents FS_{nt}^t , the fitness value $F_n(FS_{rand}^t)$ for FS_{rand}^t and $F_n(FS_{nt}^t)$ for FS_{nt}^t will be calculated to decide about the horizontal and vertical movement. In case of $F_n(FS_{rand}^t) < F_n(FS_{nt}^t)$, the movement will be vertically and it will be horizontally otherwise as follow

Case 5: Location of FS_{nt} and moving vertically or horizontally based on the fitness value $F_n(FS_{rand}^t)$:

$$FS_{nt}^{t+1} = \begin{cases} FS_{nt}^t + V_{nt}^t + \\ c_3 r(FS_{rand}^t - FS_{nt}^t) & \text{if } F_n(FS_{rand}^t) < F_n(FS_{nt}^t) \\ FS_{nt}^t + V_{nt}^t + \\ c_1 r(FS_{ht}^t - FS_{nt}^t) & \text{otherwise} \end{cases} \quad (21)$$

where c_3 is a random number $\in [0, 1]$. The last case will be applied if the condition of the horizontal and vertical movement is not achieved.

Case 6: Location of FS_{nt} and moving will be exponentially:

$$FS_{nt}^{t+1} = FS_{nt}^t + |(FS_{rand}^t - FS_{nt}^t)| \exp(bt) \cos(2\pi t) \quad (22)$$

where b is a random number $\in [0, 1]$.

The seasonal constant (S_c) and the minimal value of the seasonal constant S_{min} are calculated from Eq. 9 and Eq. 10 to check the monitoring condition ($S_c^t < S_{min}$) for t is the current iteration and t_m indicates iterations maximum value. The value of S_{min} can affect the algorithm exploration and exploitation capabilities during iterations. If specific condition is occurred, such flying squirrels's relocation is modeled

TABLE 2. Performance metrics for classification.

Metric	Value
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Specificity	$\frac{TN + FP}{TP + FP}$
Sensitivity	$\frac{TP + FN}{TN + FN}$
Negative Predictive Value (NPV)	$\frac{TN + FN}{TP + FN}$
Precision (PPV)	$\frac{TP + FP}{PPV \times TPR}$
F-score	$2 \times \frac{PPV + TPR}{PPV + TPR}$

by Eq. 23 which has the effect on local and global optima as shown in Fig. 5:

$$FS_{nt}^{new} = FS_{nt}^t + 2r((FS_{ht}^t - FS_{nt}^t)(1 - (\frac{FS_{ht}^t + FS_{nt}^t}{FS_{nt}^t})^2)) \quad (23)$$

The proposed ASSOA algorithm is explained step by step in Algorithm 2. The proposed algorithm's computational complexity will be discussed as shown in Algorithm 2. Let the number of population be $n = n_1 + n_2 + n_3$; the maximum number of iterations be t_m . For parts of the ASSOA algorithm, the time complexity will be defined as in the following points:

- Initialize of ASSOA population: $O(1)$.
- Initialize of ASSOA parameters $R_1, R_2, R_3, n_1, n_2, n_3, P_{dp}, G_c, c_1, c_2, c_3, r, b, P_a, P_d, a, d, p$: $O(1)$.
- Calculate fitness function for each agent: $O(n)$.
- Sorting agents in ascending order: $O(n)$.
- Finding first best individual, next three best individuals, normal individuals: $O(n)$.
- Positions' updating for each agent in case 1: $O(t_m \times n_1)$.
- Positions' updating for each agent in case 2: $O(t_m \times n_2)$.
- Positions' updating for each agent in case 3: $O(t_m \times n_3)$.
- Positions' updating for each agent in case 4: $O(t_m)$.
- Positions' updating for each agent in case 5: $O(t_m)$.

Algorithm 2 Proposed ASSOA Algorithm

```

1: Initialize ASSOA population  $FS_i (i = 1, 2, \dots, n)$  with size  $n$  using Eq. (15), velocities  $V_i (i = 1, 2, \dots, n)$  using Eq. (16), maximum iterations  $t_m$ , and fitness function  $F_n$ .
2: Initialize ASSOA parameters  $R_1, R_2, R_3, n_1, n_2, n_3, P_{dp}, G_c, c_1, c_2, c_3, r, b, P_a, P_d, a, d, p, t = 1$ 
3: Calculate fitness function  $F_n$  for each  $FS_i$  using Eq. (5) and Sort flying squirrels locations in ascending order
4: Find the first best individual  $FS_{ht}$ , the next three best individuals  $FS_{at}$ , the normal individuals  $FS_{nt}$ 
5: while  $t \leq t_m$  (Stopping condition) do
6:   if  $(p \geq 0.5)$  then
7:     for  $(t = 1 : n_1)$  do
8:       if  $(R_1 \geq P_{dp})$  then
9:          $FS_{at}^{t+1} = FS_{at}^t + d_g \times G_c (FS_{ht}^t - FS_{at}^t)$ 
10:      else
11:         $FS_{at}^{t+1} = \text{Random location}$ 
12:      end if
13:    end for
14:    for  $(t = 1 : n_2)$  do
15:      if  $(R_2 \geq P_{dp})$  then
16:         $FS_{nt}^{t+1} = FS_{nt}^t + d_g \times G_c (FS_{at}^t - FS_{nt}^t)$ 
17:      else
18:         $FS_{nt}^{t+1} = \text{Random location}$ 
19:      end if
20:    end for
21:    for  $(t = 1 : n_3)$  do
22:      if  $(R_3 \geq P_{dp})$  then
23:         $FS_{nt}^{t+1} = FS_{nt}^t + d_g \times G_c (FS_{ht}^t - FS_{nt}^t)$ 
24:      else
25:         $FS_{nt}^{t+1} = \text{Random location}$ 
26:      end if
27:    end for
28:  else
29:    if  $(P_a < a)$  then
30:       $FS_{nt}^{t+1} = FS_{nt}^t + V_{nt}^t + c_1 r (FS_{ht}^t - FS_{nt}^t) + c_2 r (FS_{at}^t - FS_{nt}^t)$ 
31:    else
32:      Choose random agent  $FS_{rand}^t$  from normal agents  $FS_{nt}^t$ 
33:      if  $(P_d < d)$  then
34:        Calculate fitness function  $F_n (FS_{rand}^t)$  for  $FS_{rand}^t$ 
35:        if  $(F_n (FS_{rand}^t) < F_n (FS_{nt}^t))$  then
36:           $FS_{nt}^{t+1} = FS_{nt}^t + V_{nt}^t + c_3 r (FS_{rand}^t - FS_{nt}^t)$ 
37:        else
38:           $FS_{nt}^{t+1} = FS_{nt}^t + V_{nt}^t + c_1 r (FS_{ht}^t - FS_{nt}^t)$ 
39:        end if
40:      else
41:         $FS_{nt}^{t+1} = FS_{nt}^t + |(FS_{rand}^t - FS_{nt}^t)| \exp(bt) \cos(2\pi t)$ 
42:      end if
43:    end if
44:  end if
45:  Calculate seasonal constant  $(S_c^t)$  using Eq. (9)
46:  Calculate minimum value of seasonal constant  $(S_{min})$  using Eq. (10)
47:  if  $(S_c^t < S_{min})$  then
48:     $FS_{nt}^{new} = FS_{ht}^t + 2 r ((FS_{ht}^t - FS_{nt}^t) (1 - (\frac{FS_{ht}^t + FS_{nt}^t}{FS_{nt}^t})^2))$ 
49:  end if
50:  Update  $S_{min}$  using Eq. (10)
51:  Set  $t = t + 1$ 
52: end while
53: Return optimal solution  $FS_{ht}$ 

```

- Positions' updating for each agent in case 6: $O(t_m)$.
- Calculating the seasonal constant: $O(t_m)$.
- Calculating the minimum value of seasonal constant: $O(t_m)$.
- Relocation of agents: $O(t_m)$.

- Iteration number increment : $O(t_m)$.
- Returning the best individual: $O(1)$

The above analysis indicates that the proposed ASSOA algorithm' complexity of computations is $O(t_m \times n)$ and in case of a problem with d dimension is $O(t_m \times n \times d)$.

TABLE 3. Performance metrics outputs of the compared deep learning approaches.

Method/Metric	Accuracy	Sensitivity	Specificity	PPV	NPV	F-score	Time (s)
AlexNet	0.820	0.840	0.610	0.670	0.850	0.745	724
VGGNet	0.869	0.741	0.751	0.997	0.249	0.850	747
GoogLeNet	0.880	0.910	0.890	0.800	0.920	0.851	841
ResNet-50	0.910	0.950	0.920	0.840	0.950	0.892	203

TABLE 4. CNN experimental setup.

Parameter	Value
CNN training options (Default)	
RateDropFactor	0.1000
Momentum Learn	0.9000
L2Regularization	1.0000e-04
LearnRateDropPeriod	10
GradientThreshold	Inf
GradientThresholdMethod	l2norm
ValidationData	imds
VerboseFrequency	50
ValidationPatience	Inf
ValidationFrequency	50
ResetInputNormalization	1
CNN training options (Custom)	
InitialLearnRate	1.0000e-04
ExecutionEnvironment	gpu
MiniBatchSize	8
MaxEpochs	20
Verbose	0
Shuffle	every-epoch
LearnRateSchedule	piecwise
Optimizer	sgdm

TABLE 5. Performance metrics for feature selection.

Metric	Value
Average Error	$1 - \frac{1}{M} \sum_{j=1}^M \frac{1}{N} \sum_{i=1}^N Match(C_i, L_i)$
Average Select Size	$\frac{1}{M} \sum_{j=1}^M \frac{size(g_j^*)}{D}$
Average Fitness	$\frac{1}{M} \sum_{j=1}^M g_j^*$
Best Fitness	$Min_{j=1}^M g_j^*$
Worst Fitness	$Max_{j=1}^M g_j^*$
Standard Deviation	$\sqrt{\frac{1}{M-1} \sum (g_j^* - Mean)^2}$

TABLE 6. Proposed ASSOA algorithm configuration.

Parameter	Value
Number of iterations	50
Number of agents	10
R_1	[0,1]
R_2	[0,1]
R_3	[0,1]
G_c	1.9
P_{dp}	0.1
c_1, c_2, c_3	[0,1]
P_a	[0,1]
P_d	[0,1]
a, r, b, p	[0,1]

C. BINARY OPTIMIZER

In the feature selection problem, the search space is represented by only binary values of 0 and 1. Thus, the proposed ASSOA algorithm’s continuous values are converted into binary values for the process of feature selection based on the ResNet-50 model’s extracted features. The following

TABLE 7. Compared algorithms configuration.

Algorithm	Parameter (s)	Value (s)
SS	Number of iterations	50
	Number of agents	10
	R_1	[0,1]
	R_2	[0,1]
	R_3	[0,1]
	G_c	1.9
	P_{dp}	0.1
GWO	No. of iterations	50
	No. of wolves	10
	a	2 to 0
GA	Generations	50
	Population size	10
	Mutation ratio	0.1
	Crossover	0.9
	Selection mechanism	Roulette wheel

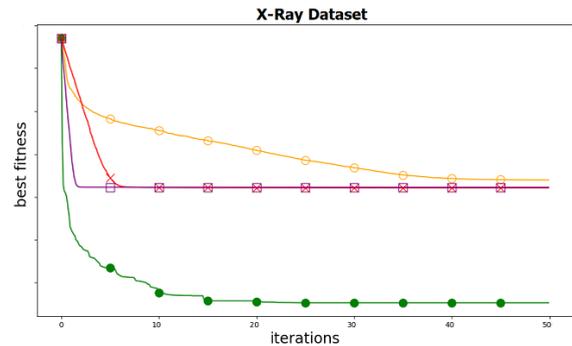


FIGURE 6. Proposed ASSOA Convergence curves compared to other techniques; Green, purple, red, and yellow lines indicates ASSOA, SS, GWO, GA algorithms, respectively.

TABLE 8. Performance of the proposed ASSOA algorithm for feature selection compared to other algorithms.

Metric / Optimizer	ASSOA	SS	GWO	GA
Average error	0.2113	0.2956	0.3243	0.3317
Average Select size	0.7333	0.7967	0.8000	0.8000
Average Fitness	0.1236	0.1602	0.1673	0.1752
Best Fitness	1.2812	1.3012	1.3909	1.4009
Worst Fitness	1.4061	1.4854	1.5057	1.5142
Standard deviation Fitness	0.0047	0.0055	0.0061	0.0067

equation will be applied to get the binary values from the standard continuous values of the proposed ASSOA algorithm.

$$FS_d^{(t+1)} = \begin{cases} 1 & \text{if } Sigmoid(x) \geq 0.5 \\ 0 & \text{otherwise,} \end{cases}$$

$$Sigmoid(x) = \frac{1}{1 + \exp^{-10(x-0.5)}}, \quad (24)$$

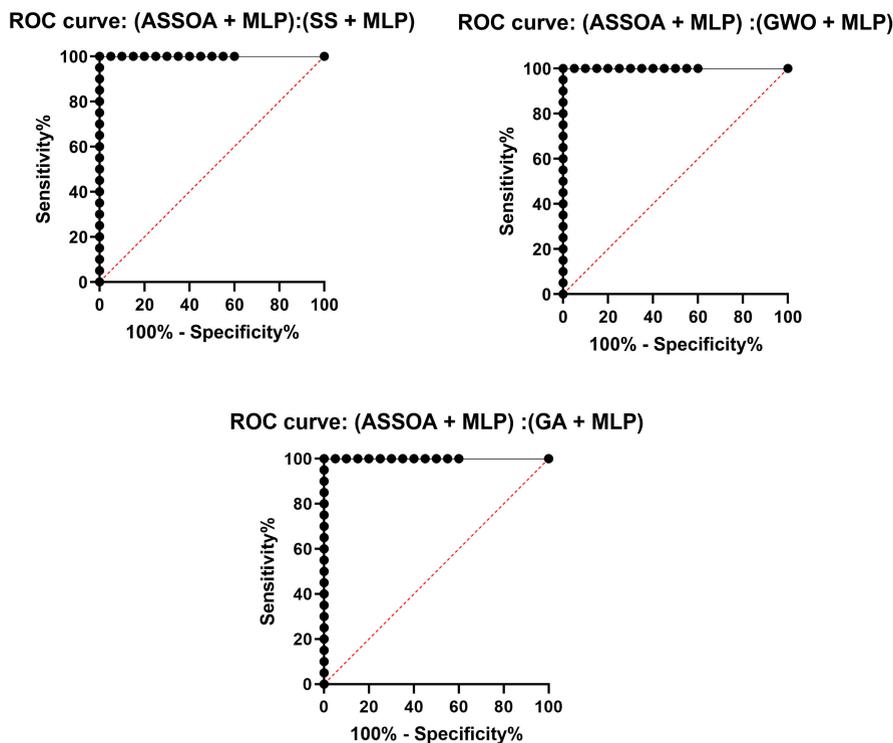


FIGURE 7. ROC curves for ASSOA algorithm versus compared algorithms.

TABLE 9. Proposed ASSOA algorithm’s p-values in comparison to other algorithms using Wilcoxon’s rank-sum based on average error metric.

Compared Algorithms	SS	GWO	GA
p-value	1.21E-05	1.21E-05	1.21E-05

where $FS_d^{(t+1)}$ is the binary position at iteration t of d dimension. The *Sigmoid* function scales the continuous values to be zero or one. $Sigmoid(x) \geq 0.5$ condition is employed here to filter the values to be 0 or 1. the x value indicates the best solution of the algorithm which is denoted as FS_{ht} in Algorithm 2.

D. FITNESS FUNCTION

The fitness function measures the optimizer solutions’ quality. The function is dependent on the classification error rate and the selected features. The excellent solution corresponds to a set of features that give lower features and classification error rate. To evaluate the solution quality, Eq. 25 can be employed

$$F_n = h_1 Err(O) + h_2 \frac{|s|}{|f|} \tag{25}$$

where $Err(O)$ indicates the optimizer error rate, s denotes the set of features selected by the optimizer, f denotes the features’ total number. The $h_1 \in [0, 1], h_2 = 1 - h_1$ values manage the importance of the error rate of classification process and the selected feature number.

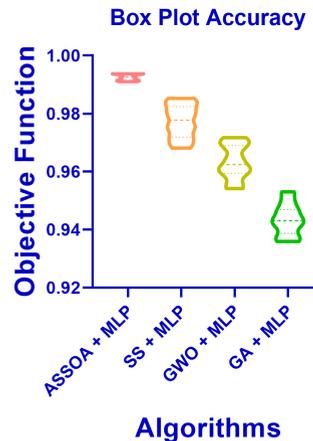


FIGURE 8. Box plot accuracy of the proposed ASSOA + MLP algorithm compared to other algorithms for chest X-ray dataset.

V. EXPERIMENTAL RESULTS

There are three scenarios in the experiments. The first scenario shows the effectiveness of four CNN models for classifying the chest X-ray cases and offers the importance of features extraction for the next stage. The second scenario is designed to test and compare the proposed ASSOA algorithm to other optimization algorithms for feature selection. The third scenario is conducted to test the proposed ASSOA algorithm’s ability as a classifier for improving the classification accuracy based on MLP. Wilcoxon’s rank-sum test

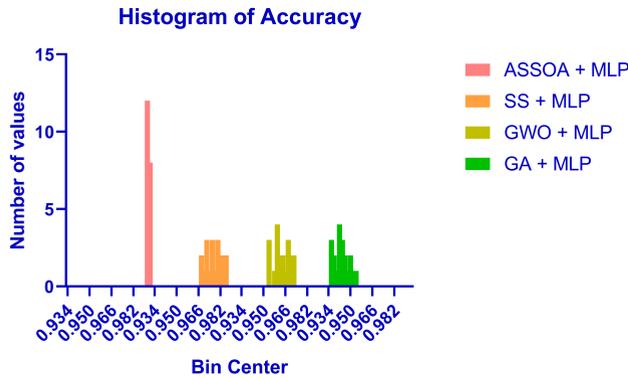


FIGURE 9. Histogram of accuracy of the proposed ASSOA + MLP algorithm compared to other algorithms for chest X-ray dataset.

is performed to verify the proposed algorithm’s superiority statistically. For the chest X-ray datasets, the images are separated randomly into training images of (60%), validation images of (20%), testing images of (20%). The data in the training process is used to train the CNN model. In contrast, the validation process data is applied for verification purposes, and the testing data evaluated the efficiency of the proposed method for the unknown chest X-ray cases.

A. FIRST SCENARIO

The classification accuracy of the four CNN models namely AlexNet [56], VGGNet [57], GoogLeNet [58], and ResNet-50 [60] is calculated in this scenario for the tested chest X-ray dataset. Let *TP* indicates true-positive value, *FP* represents false-positive value, *TN* indicates true-negative value, and *FN* represents false-negative value. The performance metrics, such as accuracy, precision, and F-score [29], are calculated to measure the classification performance of the CNN models as shown in Table 2. The results of this scenario including the required CPU time are shown in Table 3.

Table 4 presents the settings of the CNN experimental setup in this scenario. The default parameters are used in this case since the current stage is employed for feature extraction of the chest X-ray images from the CNN model to be used for the next scenario. The highest accuracy achieved in this case, for the X-ray images, is (91.0%) by the ResNet-50 model with an F-score of (89.2%) and required time of (203) seconds. According to the promising performance of the ResNet-50 model, a set of features is extracted from the model’s earlier layers since the model accuracy should be improved for the critical cases. In the second scenario, these features are employed to extract the best classification features by the proposed ASSOA algorithm.

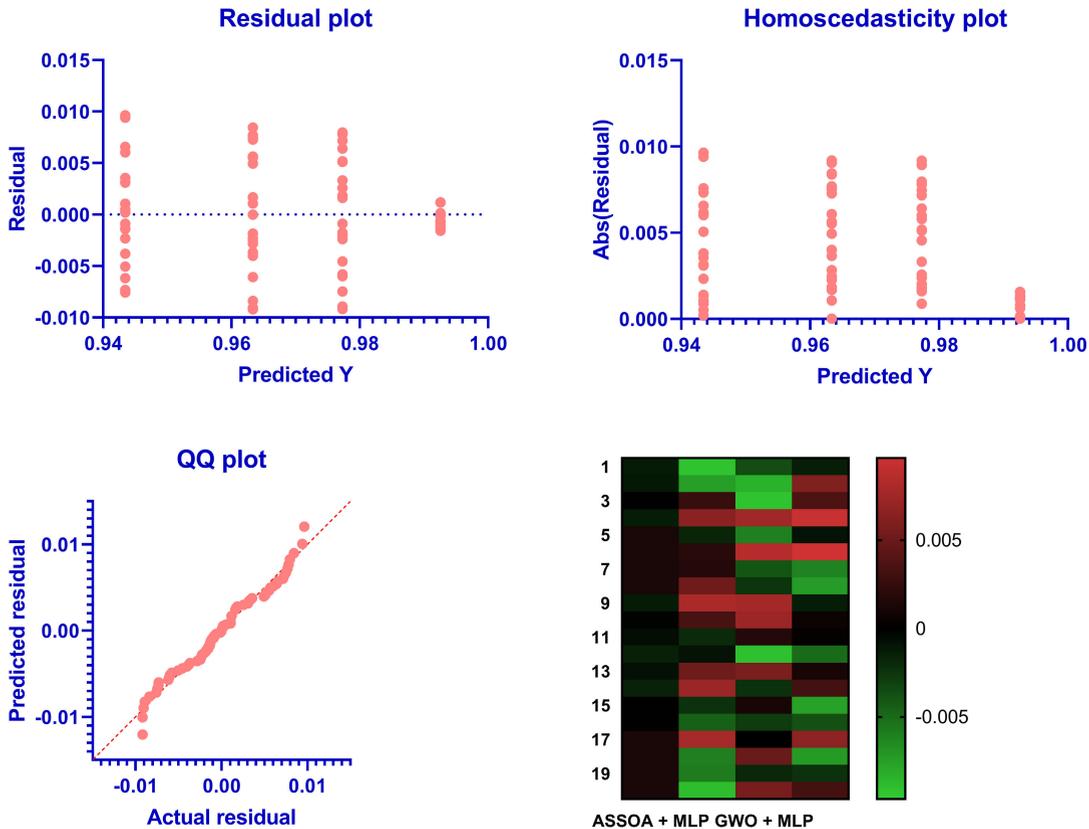


FIGURE 10. Residuals Versus Fits Plot of the proposed ASSOA + MLP algorithm compared to other algorithms for chest X-ray dataset.

TABLE 10. Descriptive statistics and the classification accuracy of the ASSOA algorithm based on MLP in comparison to other algorithms for chest X-ray images.

	ASSOA + MLP	SS + MLP	GWO + MLP	GA + MLP
Number of values	20	20	20	20
Minimum (Accuracy)	0.991	0.9681	0.9541	0.9359
25% Percentile	0.9914	0.9718	0.9594	0.9387
Median (Accuracy)	0.9926	0.9777	0.9624	0.9431
75% Percentile	0.9938	0.9825	0.969	0.9469
Maximum (Accuracy)	0.9938	0.9853	0.9718	0.9531
Range	0.002737	0.01715	0.01763	0.01722
Mean (Accuracy)	0.9926	0.9773	0.9633	0.9435
Std. Deviation	0.001098	0.0057	0.005878	0.005367
Std. Error of Mean	0.0002455	0.001275	0.001314	0.0012
Lower 95% CI of mean	0.9921	0.9746	0.9606	0.9409
Upper 95% CI of mean	0.9931	0.98	0.9661	0.946
Coefficient of variation	0.1106%	0.5832%	0.6102%	0.5688%
Geometric mean	0.9926	0.9773	0.9633	0.9434
Geometric SD factor	1.001	1.006	1.006	1.006
Lower 95% CI of geo. mean	0.9921	0.9746	0.9606	0.9409
Upper 95% CI of geo. mean	0.9931	0.98	0.9661	0.946
Harmonic mean	0.9926	0.9773	0.9633	0.9434
Lower 95% CI of harm. mean	0.9921	0.9746	0.9606	0.9409
Upper 95% CI of harm. mean	0.9931	0.9799	0.9661	0.9459
Lower 95% CI of quad. mean	0.9921	0.9746	0.9606	0.941
Upper 95% CI of quad. mean	0.9931	0.98	0.9661	0.946
Skewness	-0.1304	-0.1544	-0.04561	0.2562
Kurtosis	-1.721	-1.269	-1.248	-0.8069
Sum	19.85	19.55	19.27	18.87

B. SECOND SCENARIO

In this scenario, the efficiency of feature selection by the proposed ASSOA algorithm is investigated. ASSOA algorithm performance is compared with the basic Squirrel Search (SS) optimization algorithm [32], Grey Wolf Optimizer (GWO) [29], and Genetic Algorithm (GA) [33] based on performance metrics shown in Table 5. Let M be the number of runs of an optimizer; g_j^* represents the best solution at the run number j ; $size(g_j^*)$ is the size of the vector g_j^* . N is the number of tested points; C_i is the classifier's output label for a point i ; L_i is the class's label for a point i ; the total number of features (D); and matching between two inputs is calculated by *Match* function. The metrics used in this scenario are average error, select size, fitness, best and worst fitness, and standard deviation fitness. ASSOA algorithm configuration setting is shown in Table 6. h_1 parameter in the objective function is assigned to 0.99 and h_2 parameter to 0.01. The configuration of the SS, GWO, GA algorithms, including the number of iterations, agents, and parameters, is shown in Table 7.

Table 8 shows the ASSOA, SS, GWO, and GA algorithms' output results in this scenario. For the displayed results, if the optimizer can select the proper set of features, the error is minimized. ASSOA can achieve a minimum average error of

(0.2113) for feature selection. Based on the tested problem's minimum error, the ASSOA algorithm is the best, and GA is the worst. This means that the proposed ASSOA algorithm achieved better results than the original SS algorithm. In terms of standard deviation, the ASSOA algorithm has the lowest value than other algorithms that indicate the algorithm's robustness and stability. Figure 6 shows the ASSOA convergence curve compared to different algorithms. The figure demonstrates the optimizer exploitation capability and the ability of the algorithm to avoid local optima. The figure results show the reliability and robustness of the ASSOA algorithm to get the optimal set of features in a minimum time.

The p-values of the ASSOA algorithm are tested compared to SS, GWO, and GA algorithms by Wilcoxon's rank-sum test. The employed test can get if there is a significant difference between the ASSOA algorithm and other algorithms. If the p-value < 0.05 , this indicates that the ASSOA algorithm results are significantly different from other algorithms. If p-value > 0.05 , this indicates that the algorithm results have no significant difference. The p-value results in this scenario are shown in Table 9. Results show that the p-values are less than 0.05, proving the superiority of the proposed

TABLE 11. ANOVA test results of the proposed ASSOA + MLP algorithm for chest X-ray dataset.

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	0.02618	3	0.008728	F (3, 76) = 359.7	P < 0.0001
Residual (within columns)	0.001844	76	2.43E-05	-	-
Total	0.02803	79	-	-	-

ASSOA algorithm and that the algorithm has statistically significant.

C. THIRD SCENARIO

The last scenario checks the classification accuracy of the ASSOA algorithm based on MLP (ASSOA + MLP) in comparison with other algorithms of SS + MLP, GWO + MLP, and GA + MLP. The classification performance is tested for chest X-ray cases and other cases based on chest X-ray COVID-19. The configuration of the proposed ASSOA algorithm and the compared algorithms are shown in Table 6 and Table 7, respectively. Metrics of the classification performance used in this scenario are presented in Table 2.

1) CHEST X-RAY CLASSIFICATION RESULTS

The results of the ASSOA + MLP algorithm and other algorithms regarding accuracy are shown in Table 10. The proposed algorithm (ASSOA + MLP) from the descriptive statistics, as shown in Table 10, can achieve a mean accuracy of (99.26%) and a standard deviation of (0.001098) within (135) seconds to classify a new input X-ray chest image which outperforms other algorithms. The ROC curves of the proposed ASSOA algorithm based on MLP versus the compared classification algorithms are shown in Figure 7. From this figure, the proposed algorithm can highly distinguish among the X-ray chest images with a high AUC value equal to (0.9875). The Box plot accuracy and Histogram of accuracy are also tested, and the output figures are shown in Figures 8 and 9. These figures show the stability and consistency of the proposed algorithm for the classification of different cases.

Wilcoxon’s rank-sum and ANOVA tests are performed in this scenario to get the ASSOA + MLP algorithm’s p-values compared to SS + MLP, GWO + MLP, GA + MLP classification algorithms. These tests can indicate the significant difference between the ASSOA + MLP algorithm results and compared algorithms. The output p-values are shown in Table 12 for Wilcoxon’s rank-sum test, and in Table 11 for ANOVA test. Note that the p-values are less than 0.05, which indicates the superiority of the ASSOA + MLP algorithm and that the algorithm is statistically significant.

The possible problems can be observed from the residual values, and residual plots rather than the original dataset plot since some datasets are not good candidates for classification. The ideal case is achieved if the residual values are equally and randomly spaced around the horizontal axis. The residual value can be calculated as (Actual value - Predicted value) with the mean and sum of the residuals are equal to zero. A residual plot is used to present the vertical axis’s residual values and the independent variable on the horizontal axis.

TABLE 12. The ASSOA + MLP algorithm’s p-values in comparison to other algorithms using Wilcoxon’s rank-sum based on accuracy metric.

Compared Algorithms	SS + MLP	GWO + MLP	GA + MLP
p-value	1.11E-05	1.11E-05	1.11E-05

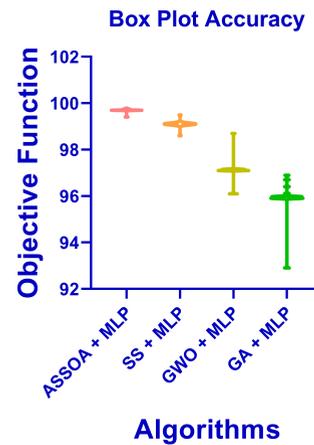


FIGURE 11. Box plot accuracy of the proposed ASSOA + MLP algorithm compared to other algorithms for chest X-ray COVID-19 dataset.

Figure 10 shows the residual plot. A linear or a nonlinear model can be decided from plot patterns in a residual plot, and an appropriate one can be determined.

The homogeneity of variance (heteroscedasticity) provides a visual examination between the prediction errors and the predicted dependent variable scores. The heteroscedasticity plot, shown in Figure 12, can quickly determine any violation and easily improve the research findings’ accuracy. Homoscedasticity describes a situation in which the error term (random disturbance in the relationship between the dependent variable and the independent variables, or noise) is the same across the independent variables’ values.

The quantile-quantile (QQ) plot is also shown in Figure 12. It is known as a probability plot. It is mainly used by plotting their quantiles against each other to compare two probability distributions. It is noted from the figure that the points distributions in the QQ plot are approximately fit on the line. Thus, the actual and the predicted residuals are linearly related, confirming the proposed ASSOA + MLP classifier’s performance to identify the chest X-ray images.

2) CHEST X-RAY COVID-19 CLASSIFICATION RESULTS

A chest X-ray COVID-19 dataset [31] is tested in the experiments to test the performance of the proposed ASSOA + MLP algorithm for the classification of chest X-ray

TABLE 13. Descriptive statistics and the classification accuracy of the ASSOA algorithm based on MLP in comparison to other algorithms for chest X-ray COVID-19 images.

	ASSOA + MLP	SS + MLP	GWO + MLP	GA + MLP
Number of values	20	20	20	20
Minimum (Accuracy)	0.994	0.986	0.961	0.929
25% Percentile	0.997	0.991	0.971	0.959
Median (Accuracy)	0.997	0.991	0.9712	0.9596
75% Percentile	0.9974	0.991	0.9715	0.9598
Maximum (Accuracy)	0.998	0.995	0.987	0.969
Range	0.004	0.009	0.026	0.04
Mean (Accuracy)	0.997	0.991	0.971	0.959
Std. Deviation	0.0007615	0.001478	0.00491	0.007597
Std. Error of Mean	0.0001703	0.0003305	0.001098	0.001699
Lower 95% CI of mean	0.9967	0.9903	0.9687	0.9555
Upper 95% CI of mean	0.9974	0.9917	0.9733	0.9626
Coefficient of variation	0.07638%	0.1492%	0.5057%	0.7922%
Geometric mean	0.997	0.991	0.971	0.959
Geometric SD factor	1.001	1.001	1.005	1.008
Lower 95% CI of geo. mean	0.9967	0.9903	0.9687	0.9554
Upper 95% CI of geo. mean	0.9974	0.9917	0.9733	0.9626
Harmonic mean	0.997	0.991	0.971	0.959
Lower 95% CI of harm. mean	0.9967	0.9903	0.9687	0.9553
Upper 95% CI of harm. mean	0.9974	0.9917	0.9733	0.9626
Lower 95% CI of quad. mean	0.9967	0.9903	0.9687	0.9555
Upper 95% CI of quad. mean	0.9974	0.9917	0.9733	0.9625
Skewness	-3.421	-1.099	1.029	-3.388
Kurtosis	14.35	9.79	7.08	14.38
Sum	19.94	19.82	19.42	19.18

TABLE 14. ANOVA test results of the proposed algorithm for chest X-ray COVID-19 dataset.

ANOVA table	SS	DF	MS	F (DFn, DFd)	P value
Treatment (between columns)	0.01861	3	0.006204	F (3, 76) = 293.4	P < 0.0001
Residual (within columns)	0.001607	76	2.12E-05	-	-
Total	0.02022	79	-	-	-

COVID-19 cases. The output descriptive statistics of this experiment are shown in Table 13. The proposed (ASSOA + MLP) algorithm achieved a mean accuracy of (99.7%) for the COVID-19 dataset. The mean accuracy of compared algorithms of SS + MLP, GWO + MLP, and GA + MLP are (99.1%), (97.1%), and (95.9%), respectively. These results show that the proposed algorithm can improve the classification accuracy of COVID-19 patients from their chest X-ray images. The Box plot accuracy is tested, and the output figure is shown in Fig. 11. This figure shows the stability and consistency of the proposed algorithm for the classification of COVID-19 cases.

ANOVA test is also performed for this experiment to test the ASSOA + MLP algorithm's p-values compared to SS + MLP, GWO + MLP, GA + MLP classification algorithms.

The output p-values are shown in Table 11 for the ANOVA test. Note that the p-values are less than 0.05, which indicates the superiority of the ASSOA + MLP algorithm and that the algorithm is statistically significant.

VI. DISCUSSION

The experiments are divided into three different scenarios to assess the proposed method performance to classify chest X-ray images. According to the promising performance, the first scenario shows that the features can be extracted from the earlier layers of the ResNet-50 model. The extracted features are fed to the next scenario for feature selection. The second scenario shows the robustness and reliability of the ASSOA algorithm in finding the optimal subset of features in a reasonable amount of time. In this scenario,

Wilcoxon's rank-sum test emphasizes the superiority of the proposed ASSOA algorithm and shows that the algorithm is statistically significant. In the third scenario, the experiments show that the proposed algorithm (ASSOA + MLP) can achieve a mean accuracy of (99.26%) and an AUC value equal to (0.9875) within (135) seconds to classify a new input X-ray chest image which outperforms other algorithms. The ASSOA + MLP algorithm also achieved a classification mean accuracy of (99.7%) for a chest X-ray COVID-19 dataset. Wilcoxon's rank-sum and ANOVA tests confirm the proposed algorithm's superiority and that the algorithm is statistically significant. The results and statistical tests demonstrate the high effectiveness of the proposed method in determining the infected cases.

VII. CONCLUSION

Developing a classification model for diagnosing infected cases is considered one of the most critical problems, which is still much too pricey for the mass selection. This paper proposes a classification model to detect the afflicted instances from the chest X-ray images, which may dramatically minimize the diagnosis prices, particularly in cultivating nations. The training and feature extraction processes are based on a convolutional neural network (CNN) based model (ResNet50) with fine-tuning and image augmentation. The X-ray images' classification to viral, normal, and bacterial, and popular scenarios are based upon an MLP neural network along with the proposed ASSOA algorithm. In this work, the chest X-ray images (Pneumonia) dataset composed of 5,863 X-ray images are utilized in the experiments. In the proposed model, a transfer learning technique is applied during the training stage and feature extraction. Experimental results show the proposed classification model's efficiency in classifying the afflicted situations and a mean accuracy of (99.26%), which surpasses the cutting-edge strategies discovered in the literature. The proposed (ASSOA + MLP) algorithm also achieved a classification mean accuracy of (99.7%) for another chest X-ray COVID-19 dataset.

ACKNOWLEDGMENT

The authors would like to thank Dr. Mohamed Elsayed Gawish and Dr. Shaaban Omar for their support in interpreting the X-ray datasets. and separate X-ray images of the affected patients from other non-infected

REFERENCES

- [1] A. Ibrahim and E.-S. M. El-Kenawy, "Image segmentation methods based on superpixel techniques: A survey," *J. Comput. Sci. Inf. Syst.*, vol. 15, no. 3, pp. 1–11, 2020. [Online]. Available: <http://www.jcsis.org/papers/15-46>
- [2] C. Li, D. Zhang, S. Du, and Z. Tian, "Deformation and refined features based lesion detection on chest X-ray," *IEEE Access*, vol. 8, pp. 14675–14689, 2020, doi: [10.1109/access.2020.2963926](https://doi.org/10.1109/access.2020.2963926).
- [3] A. Ibrahim, S. Mohammed, H. A. Ali, and S. E. Hussein, "Breast cancer segmentation from thermal images based on chaotic salp swarm algorithm," *IEEE Access*, vol. 8, pp. 122121–122134, 2020, doi: [10.1109/access.2020.3007336](https://doi.org/10.1109/access.2020.3007336).
- [4] A. Ibrahim, S. Mohammed, and H. A. Ali, "Breast cancer detection and classification using thermography: A review," in *Proc. Int. Conf. Adv. Mach. Learn. Technol. Appl. (AMLTA)*, A. E. Hassanien, M. F. Tolba, M. Elhoseny, and M. Mostafa, Eds. Cham, Switzerland: Springer, 2018, pp. 496–505.
- [5] S. Xu, H. Wu, and R. Bie, "CXNet-m1: Anomaly detection on chest X-rays with image-based deep learning," *IEEE Access*, vol. 7, pp. 4466–4477, 2019, doi: [10.1109/access.2018.2885997](https://doi.org/10.1109/access.2018.2885997).
- [6] E.-S. M. El-Kenawy and M. Eid, "Hybrid gray wolf and particle swarm optimization for feature selection," *Int. J. Innov. Comput. Inf. Control*, vol. 16, no. 3, pp. 831–844, 2020.
- [7] A. Ibrahim, T. Gaber, T. Horiuchi, V. Snaesl, and A. E. Hassanien, "Human thermal face extraction based on superpixel technique," in *Proc. 1st Int. Conf. Adv. Intell. Syst. Inform. (AISII)*, E. B. Suf, T. Gaber, A. E. Hassanien, N. El-Bendary, and N. Dey, Eds. Cham, Switzerland: Springer, 2016, pp. 163–172.
- [8] T. Gaber, A. Tharwat, A. Ibrahim, V. Snael, and A. E. Hassanien, "Human thermal face recognition based on random linear oracle (RLO) ensembles," in *Proc. Int. Conf. Intell. Netw. Collaborative Syst.*, Sep. 2015, pp. 91–98, doi: [10.1109/incos.2015.67](https://doi.org/10.1109/incos.2015.67).
- [9] G. Shi, H. Huang, and L. Wang, "Unsupervised dimensionality reduction for hyperspectral imagery via local geometric structure feature learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 8, pp. 1425–1429, Aug. 2020, doi: [10.1109/lgrs.2019.2944970](https://doi.org/10.1109/lgrs.2019.2944970).
- [10] F. Luo, B. Du, L. Zhang, L. Zhang, and D. Tao, "Feature learning using spatial-spectral hypergraph discriminant analysis for hyperspectral image," *IEEE Trans. Cybern.*, vol. 49, no. 7, pp. 2406–2419, Jul. 2019, doi: [10.1109/tcyb.2018.2810806](https://doi.org/10.1109/tcyb.2018.2810806).
- [11] R. Hang and Q. Liu, "Dimensionality reduction of hyperspectral image using spatial regularized local graph discriminant embedding," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 9, pp. 3262–3271, Sep. 2018, doi: [10.1109/jstars.2018.2847042](https://doi.org/10.1109/jstars.2018.2847042).
- [12] Y. Dong, B. Du, L. Zhang, and L. Zhang, "Dimensionality reduction and classification of hyperspectral images using ensemble discriminative local metric learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2509–2524, May 2017, doi: [10.1109/tgrs.2016.2645703](https://doi.org/10.1109/tgrs.2016.2645703).
- [13] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "DeTrac: Transfer learning of class decomposed medical images in convolutional neural networks," *IEEE Access*, vol. 8, pp. 74901–74913, 2020, doi: [10.1109/access.2020.2989273](https://doi.org/10.1109/access.2020.2989273).
- [14] G. Choy, O. Khalilzadeh, M. Michalski, S. Do, A. E. Samir, O. S. Pianykh, J. R. Geis, P. V. Pandharipande, J. A. Brink, and K. J. Dreyer, "Current applications and future impact of machine learning in radiology," *Radiology*, vol. 288, no. 2, pp. 318–328, Aug. 2018, doi: [10.1148/radiol.2018171820](https://doi.org/10.1148/radiol.2018171820).
- [15] J. A. Dunnmon, D. Yi, C. P. Langlotz, C. Ré, D. L. Rubin, and M. P. Lungren, "Assessment of convolutional neural networks for automated classification of chest radiographs," *Radiology*, vol. 290, no. 2, pp. 537–544, Feb. 2019, doi: [10.1148/radiol.2018181422](https://doi.org/10.1148/radiol.2018181422).
- [16] E.-S. M. El-Kenawy, A. Ibrahim, S. Mirjalili, M. M. Eid, and S. E. Hussein, "Novel feature selection and voting classifier algorithms for COVID-19 classification in CT images," *IEEE Access*, vol. 8, pp. 179317–179335, 2020, doi: [10.1109/access.2020.3028012](https://doi.org/10.1109/access.2020.3028012).
- [17] E.-S. M. El-Kenawy, M. M. Eid, and A. Ibrahim, "An advanced patient health monitoring system," *J. Comput. Sci. Inf. Syst.*, vol. 17, no. 11, pp. 1–7, 2021. [Online]. Available: <http://www.jcsis.org/papers/17-49>
- [18] E.-S. M. El-Kenawy, M. M. Eid, and A. Ibrahim, "Anemia estimation for COVID-19 patients using a machine learning model," *J. Comput. Sci. Inf. Syst.*, vol. 17, no. 11, pp. 1–7, 2021. [Online]. Available: <http://www.jcsis.org/papers/17-51>
- [19] M. Chumbita, C. Cillóniz, P. Puerta-Alcalde, E. Moreno-García, G. Sanjuan, N. Garcia-Pouton, A. Soriano, A. Torres, and C. Garcia-Vidal, "Can artificial intelligence improve the management of pneumonia," *J. Clin. Med.*, vol. 9, no. 1, p. 248, Jan. 2020, doi: [10.3390/jcm9010248](https://doi.org/10.3390/jcm9010248).
- [20] V. Chouhan, S. K. Singh, A. Khamparia, D. Gupta, P. Tiwari, C. Moreira, R. Damaševičius, and V. H. C. de Albuquerque, "A novel transfer learning based approach for pneumonia detection in chest X-ray images," *Appl. Sci.*, vol. 10, no. 2, p. 559, Jan. 2020, doi: [10.3390/app10020559](https://doi.org/10.3390/app10020559).
- [21] N. Bacanin, T. Bezdan, E. Tuba, I. Strumberger, and M. Tuba, "Optimizing convolutional neural network hyperparameters by enhanced swarm intelligence metaheuristics," *Algorithms*, vol. 13, no. 3, p. 67, Mar. 2020, doi: [10.3390/a13030067](https://doi.org/10.3390/a13030067).

- [22] A. Banharsakun, "Towards improving the convolutional neural networks for deep learning using the distributed artificial bee colony method," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 6, pp. 1301–1311, Jun. 2019.
- [23] M. M. Eid and Y. H. Elawady, "Efficient pneumonia detection for chest radiography using ResNet-based SVM," *Eur. J. Electr. Eng. Comput. Sci.*, vol. 5, no. 1, pp. 1–8, Jan. 2021, doi: [10.24018/ejece.2021.5.1.268](https://doi.org/10.24018/ejece.2021.5.1.268).
- [24] M. M. Fouad, A. I. El-Desouky, R. Al-Hajj, and E.-S. M. El-Kenawy, "Dynamic group-based cooperative optimization algorithm," *IEEE Access*, vol. 8, pp. 148378–148403, 2020, doi: [10.1109/access.2020.3015892](https://doi.org/10.1109/access.2020.3015892).
- [25] A. Ibrahim, A. Tharwat, T. Gaber, and A. E. Hassanien, "Optimized superpixel and AdaBoost classifier for human thermal face recognition," *Signal, Image Video Process.*, vol. 12, no. 4, pp. 711–719, May 2018, doi: [10.1007/s11760-017-1212-6](https://doi.org/10.1007/s11760-017-1212-6).
- [26] E. M. Hassib, A. I. El-Desouky, E.-S. M. El-Kenawy, and S. M. El-Ghamrawy, "An imbalanced big data mining framework for improving optimization algorithms performance," *IEEE Access*, vol. 7, pp. 170774–170795, 2019.
- [27] A. Ibrahim, M. Noshay, H. A. Ali, and M. Badawy, "PAPSO: A power-aware VM placement technique based on particle swarm optimization," *IEEE Access*, vol. 8, pp. 81747–81764, 2020, doi: [10.1109/access.2020.2990828](https://doi.org/10.1109/access.2020.2990828).
- [28] E. M. Hassib, A. I. El-Desouky, L. M. Labib, and E.-S. M. El-Kenawy, "WOA + BRNN: An imbalanced big data classification framework using whale optimization and deep neural network," *Soft Comput.*, vol. 24, no. 8, pp. 5573–5592, Mar. 2019, doi: [10.1007/s00500-019-03901-y](https://doi.org/10.1007/s00500-019-03901-y).
- [29] E.-S. M. El-Kenawy, M. M. Eid, M. Saber, and A. Ibrahim, "MbGWO-SFS: Modified binary grey wolf optimizer based on stochastic fractal search for feature selection," *IEEE Access*, vol. 8, no. 1, pp. 107635–107649, 2020, doi: [10.1109/access.2020.3001151](https://doi.org/10.1109/access.2020.3001151).
- [30] D. Kermary et al., "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018, doi: [10.1016/j.cell.2018.02.010](https://doi.org/10.1016/j.cell.2018.02.010).
- [31] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "COVID-19 image data collection: Prospective predictions are the future," Oct. 2020, *arXiv:2006.11988*. [Online]. Available: <https://arxiv.org/abs/2006.11988> and <https://github.com/ieee8023/covid-chestxray-dataset>
- [32] M. Jain, V. Singh, and A. Rani, "A novel nature-inspired algorithm for optimization: Squirrel search algorithm," *Swarm Evol. Comput.*, vol. 44, pp. 148–175, Feb. 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2210650217305229>
- [33] M. M. Kabir, M. Shahjahan, and K. Murase, "A new local search based hybrid genetic algorithm for feature selection," *Neurocomputing*, vol. 74, no. 17, pp. 2914–2928, Oct. 2011. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S09252312110002748>
- [34] O. Stephen, M. Sain, U. J. Maduh, and D.-U. Jeong, "An efficient deep learning approach to pneumonia classification in healthcare," *J. Healthcare Eng.*, vol. 2019, pp. 1–7, Mar. 2019, doi: [10.1155/2019/4180949](https://doi.org/10.1155/2019/4180949).
- [35] M. Grewal, T. M. Deist, J. Wiersma, P. A. N. Bosman, and T. Alderliesten, "An end-to-end deep learning approach for landmark detection and matching in medical images," 2020, *arXiv:2001.07434*. [Online]. Available: <https://arxiv.org/abs/2001.07434>
- [36] S. Sahoo, M. Dash, S. Behera, and S. Sabut, "Machine learning approach to detect cardiac arrhythmias in ecg signals: A survey," *IRBM*, vol. 41, no. 4, pp. 185–194, 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1959031819301654>
- [37] M. Grewal, M. M. Srivastava, P. Kumar, and S. Varadarajan, "RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 281–284.
- [38] R. Raman, S. Srinivasan, S. Virmani, S. Sivaprasad, C. Rao, and R. Rajalakshmi, "Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy," *Eye*, vol. 33, no. 1, pp. 97–109, Jan. 2019.
- [39] G. H. Groeneveld, R. J. van de Peppel, M. W. M. de Waal, T. J. M. Verheij, and J. T. van Dissel, "Clinical factors, C-reactive protein point of care test and chest X-ray in patients with pneumonia: A survey in primary care," *Eur. J. Gen. Pract.*, vol. 25, no. 4, pp. 229–235, Oct. 2019.
- [40] I. Sirazitdinov, M. Kholiavchenko, T. Mustafaev, Y. Yixuan, R. Kuleev, and B. Ibragimov, "Deep neural network ensemble for pneumonia localization from a large-scale chest X-ray database," *Comput. Electr. Eng.*, vol. 78, pp. 388–399, Sep. 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0045790618332890>
- [41] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, "Identifying pneumonia in chest X-rays: A deep learning approach," *Measurement*, vol. 145, pp. 511–518, Oct. 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0263224119305202>
- [42] Z. Huang, Q. Li, T. Zhang, N. Sang, and H. Hong, "Iterative weighted sparse representation for X-ray cardiovascular angiogram image denoising over learned dictionary," *IET Image Process.*, vol. 12, no. 2, pp. 254–261, Feb. 2018, doi: [10.1049/iet-ipr.2017.0518](https://doi.org/10.1049/iet-ipr.2017.0518).
- [43] Z. Huang, Y. Zhang, Q. Li, T. Zhang, and N. Sang, "Spatially adaptive denoising for X-ray cardiovascular angiogram images," *Biomed. Signal Process. Control*, vol. 40, pp. 131–139, Feb. 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1746809417302288>
- [44] C. Sezgin, M. Gunalp, S. Genc, N. Acar, E. Ustuner, A. B. Oguz, A. K. Tanriverdi, A. Demirkan, and O. Polat, "Diagnostic value of bedside lung ultrasonography in pneumonia," *Ultrasound Med. Biol.*, vol. 46, no. 5, pp. 1189–1196, May 2020.
- [45] C. Yan, R. Hui, Z. Lijuan, and Y. Zhou, "Lung ultrasound vs. chest X-ray in children with suspected pneumonia confirmed by chest computed tomography: A retrospective cohort study," *Exp. Therapeutic Med.*, vol. 19, no. 2, pp. 1363–1369, Dec. 2019.
- [46] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, "Pneumonia detection using CNN based feature extraction," in *Proc. IEEE Int. Conf. Electr., Comput. Commun. Technol. (ICEECT)*, Feb. 2019, pp. 1–7.
- [47] P. Chhikara, P. Singh, P. Gupta, and T. Bhatia, "Deep convolutional neural network with transfer learning for detecting pneumonia on chest X-rays," in *Advances in Bioinformatics, Multimedia, and Electronics Circuits and Signals*, L. C. Jain, M. Virvou, V. Piuri, and V. E. Balas, Eds. Singapore: Springer, 2020, pp. 155–168.
- [48] K. Zhang et al., "Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography," *Cell*, vol. 181, no. 6, pp. 1423.e11–1433.e11, Jun. 2020, doi: [10.1016/j.cell.2020.04.045](https://doi.org/10.1016/j.cell.2020.04.045).
- [49] C. Butt, J. Gill, D. Chun, and B. A. Babu, "Deep learning system to screen coronavirus disease 2019 pneumonia," *Appl. Intell.*, pp. 1–7, Apr. 2020, doi: [10.1007/s10489-020-01714-3](https://doi.org/10.1007/s10489-020-01714-3).
- [50] H. Panwar, P. K. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, "Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet," *Chaos, Solitons Fractals*, vol. 138, Sep. 2020, Art. no. 109944. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S096007792030343X>
- [51] M. Nour, Z. Cömert, and K. Polat, "A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization," *Appl. Soft Comput.*, vol. 97, Dec. 2020, Art. no. 106580, doi: [10.1016/j.asoc.2020.106580](https://doi.org/10.1016/j.asoc.2020.106580).
- [52] S. Hu, Y. Gao, Z. Niu, Y. Jiang, L. Li, X. Xiao, M. Wang, E. F. Fang, W. Menpes-Smith, J. Xia, H. Ye, and G. Yang, "Weakly supervised deep learning for COVID-19 infection detection and classification from CT images," *IEEE Access*, vol. 8, pp. 118869–118883, 2020, doi: [10.1109/access.2020.3005510](https://doi.org/10.1109/access.2020.3005510).
- [53] M. A. Elaziz, K. M. Hosny, A. Salah, M. M. Darwish, S. Lu, and A. T. Sahlol, "New machine learning method for image-based diagnosis of COVID-19," *PLoS ONE*, vol. 15, no. 6, Jun. 2020, Art. no. e0235187, doi: [10.1371/journal.pone.0235187](https://doi.org/10.1371/journal.pone.0235187).
- [54] A. Ibrahim and E.-S. M. El-Kenawy, "Applications and datasets for superpixel techniques: A survey," *J. Comput. Sci. Inf. Syst.*, vol. 15, no. 3, pp. 1–6, 2020. [Online]. Available: <http://www.jcsis.org/papers/15-47>
- [55] Mathworks. *Convolutional Neural Network*. Accessed: Apr. 4, 2020. [Online]. Available: <https://ch.mathworks.com/solutions/deep-learning/convolutional-neural-network.html>
- [56] J. Han, D. Zhang, G. Cheng, N. Liu, and D. Xu, "Advanced deep-learning techniques for salient and category-specific object detection: A survey," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 84–100, Jan. 2018.
- [57] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <https://arxiv.org/abs/1409.1556>
- [58] A. Al-Dhamari, R. Sudirman, and N. H. Mahmood, "Transfer deep learning along with binary support vector machine for abnormal behavior detection," *IEEE Access*, vol. 8, pp. 61085–61095, 2020.
- [59] Shalu and A. Kamboj, "A color-based approach for melanoma skin cancer detection," in *Proc. 1st Int. Conf. Secure Cyber Comput. Commun. (ICSCCC)*, 2018, pp. 508–513.

- [60] S. Yu, L. Xie, L. Liu, and D. Xia, "Learning long-term temporal features with deep neural networks for human action recognition," *IEEE Access*, vol. 8, pp. 1840–1850, 2020.
- [61] T. Bezdán, E. Tuba, I. Strumberger, N. Bacanin, and M. Tuba, "Automatically designing convolutional neural network architecture with artificial flora algorithm," in *ICT Systems and Sustainability*, M. Tuba, S. Akashe, and A. Joshi, Eds. Singapore: Springer, 2020, pp. 371–378.
- [62] P.-C. Shih, H. Li, X. Lin, and Y. Zhang, "Application of metaheuristic algorithms for determining the structure of a convolutional neural network with a small dataset," in *Proc. 3rd Int. Conf. Comput. Sci. Appl. Eng. (CSAE)*, 2019, pp. 1–7.
- [63] Y. Wang, H. Zhang, and G. Zhang, "CPSO-CNN: An efficient PSO-based algorithm for fine-tuning hyper-parameters of convolutional neural networks," *Swarm Evol. Comput.*, vol. 49, pp. 114–123, Sep. 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2210650218310083>
- [64] J. Jiang, F. Han, Q. Ling, J. Wang, T. Li, and H. Han, "Efficient network architecture search via multiobjective particle swarm optimization based on decomposition," *Neural Netw.*, vol. 123, pp. 305–316, Mar. 2020. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0893608019303971>
- [65] S. Mukherjee, A. Adhikari, and M. Roy, "Malignant melanoma detection using multi layer perceptron with optimized network parameter selection by PSO," in *Contemporary Advances in Innovative and Applicable Information Technology*, J. K. Mandal, D. Sinha, and J. Bandopadhyay, Eds. Singapore: Springer, 2019, pp. 101–109.
- [66] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.



SEYEDALI MIRJALILI (Senior Member, IEEE) is currently the Director of the Centre for Artificial Intelligence Research and Optimization, Torrens University Australia, Brisbane, QLD, USA. He is internationally recognized for his advances in swarm intelligence and optimization, including the first set of algorithms from a synthetic intelligence standpoint—a radical departure from how natural systems are typically understood—and a systematic design framework to reliably benchmark, evaluate, and propose computationally cheap robust optimization algorithms. He has published over 200 publications with over 27 000 citations and an H-index of 58. As the most cited researcher in robust optimization, he is in the list of 1% highly cited researchers and named as one of the most influential researchers in the world by the Web of Science. He is working on the applications of multi-objective and robust meta-heuristic optimization techniques as well. His research interests include robust optimization, engineering optimization, multi-objective optimization, swarm intelligence, evolutionary algorithms, and artificial neural networks. He is an Associate Editor of several journals, including *Neurocomputing*, *Applied Soft Computing*, *Advances in Engineering Software*, *Applied Intelligence*, *PLOS One*, and *IEEE ACCESS*.



ABDELHAMEED IBRAHIM (Member, IEEE) received the bachelor's and master's degrees in engineering from the Computer Engineering and Systems Department, in 2001 and 2005, respectively, and the Ph.D. degree in engineering from the Faculty of Engineering, Chiba University, Japan, in 2011. From 2001 to 2007, he was with the Faculty of Engineering, Mansoura University, Egypt, where he is currently an Associate Professor of computer engineering. He has published over 50 publications with over 750 citations and an H-index of 17. His research interests include pattern recognition, optimization, machine learning, virtualization, and live virtual machine migration. He serves as a Reviewer for the *Journal of Electronic Imaging*, *Optical Engineering*, the *IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS*, *IEEE ACCESS*, *Computer Standards and Interfaces*, *Journal of Healthcare Engineering*, *IET Image Processing*, *Multimedia Tools and Applications*, and other respected journals.



MOHAMMED ALRAHMAWY received the B.E. degree in electronics engineering from the University of Mansoura, Egypt, in 1997, the M.Sc. degree in automatic control engineering from Mansoura University, in 2001, and the Ph.D. degree in computer science from the University of York, U.K., in 2011. He was awarded a fully-funded scholarship by the Egyptian government to join the Real-time Systems Research Group, University of York, as a Ph.D. Research Student, in 2005. He joined the Department of Computer Science, Mansoura University, as a Lecturer, in 2011, where he became an Associate Professor, in 2017. His current research interests include realtime systems and languages, cloud computing, distributed and parallel computing, soft computing, image processing, computer vision, the IoT, and big data. He was the recipient of the best M.Sc. thesis award from Mansoura University in 2003.



EL-SAYED M. EL-KENAWY (Member, IEEE) is currently an Assistant Professor with the Delta Higher Institute for Engineering and Technology (DHIET), Mansoura, Egypt. Inspiring and motivating students by providing a thorough understanding of a variety of computer concepts. He has published over 25 publications with over 550 citations and an H-index of 17. He has pioneered and launched independent research programs. Adept at sometimes explaining complex concepts in an easy-to-understand manner. His research interests include optimization, artificial intelligence, machine learning, deep learning, data science, and digital marketing. He serves as a Reviewer for the journal *IEEE ACCESS*.



M. EL-SAID received the B.Sc. (Hons.), M.Sc., and Ph.D. degrees in electric power engineering from Mansoura University, Egypt, in 1981, 1987, and 1992, respectively. Since 2005, he has been a Professor of electric power systems with the Department of Electrical Engineering, Faculty of Engineering, Mansoura University. He had held several administrative positions. From 2010 to 2011, he was the Vice-Dean for Community Services and Environmental Development, Faculty of Engineering, Mansoura University. From 2014 to 2016, he was also Director of the Communication and Information Technology Centre (CITC), Mansoura University. From 2016 to 2018, he was the Dean of the Faculty of Engineering, Mansoura University. He is currently the Dean of the Delta Higher Institute of Engineering and Technology, Ministry of Higher Education, Egypt. His main research interests include power system operation and control, smart grids, and renewable energy. He is a Regular Reviewer of many scientific journals, such as the IEEE TRANSACTIONS ON POWER SYSTEMS AND POWER DELIVERY, *Electric Power Components and Systems*. He is also an Official Reviewer for Egyptian Universities Promotion Committees - Supreme Council of Universities in Egypt.



ROKAIA M. ZAKI received the master's degree from Benha University, Egypt, in 2007. She is currently an Assistant Professor with the Shoubra Faculty of Engineering, Benha University. Her research interests include mobile communications, wireless networks, and cognitive radio.



MARWA METWALLY EID received the Ph.D. degree in electronics and communications engineering from the Faculty of Engineering, Mansoura University, Egypt, in 2015. Since 2011, she has been working as an Assistant Professor with the Delta Higher Institute for Engineering and Technology. Her current research interests include image processing, encryption, wireless communication systems, and field programmable gate array (FPGA) applications.

...