Applied Soft Computing 115 (2022) 108250

Contents lists available at ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Multi-COVID-Net: Multi-objective optimized network for COVID-19 diagnosis from chest X-ray images

Tripti Goel [a], R. Murugan [a,*], Seyedali Mirjalili [b,c], Deba Kumar Chakrabartty [d]

[a] Bio-Medical Imaging Laboratory (BIOMIL), Department of Electronics and Communication Engineering, National Institute of Technology Silchar, Assam, 788010, India
[b] Centre for Artificial Intelligence Research and Optimisation; Torrens University, Fortitude Valley, Brisbane, 4006 QLD, Australia
[c] YFL (Yonsei Frontier Lab), Yonsei University, Seoul, Korea
[d] Department of Radiology, Silchar Medical College and Hospital, Assam, 788014, India

ARTICLE INFO

Article history:
Received 13 May 2021
Received in revised form 12 October 2021
Accepted 23 November 2021
Available online 9 December 2021

Keywords:
Chest X-ray images
COVID-19
Deep learning
CNN
MOGOA
Multi-objective optimization

ABSTRACT

Coronavirus Disease 2019 (COVID-19) had already spread worldwide, and healthcare services have become limited in many countries. Efficient screening of hospitalized individuals is vital in the struggle toward COVID-19 through chest radiography, which is one of the important assessment strategies. This allows researchers to understand medical information in terms of chest X-ray (CXR) images and evaluate relevant irregularities, which may result in a fully automated identification of the disease. Due to the rapid growth of cases every day, a relatively small number of COVID-19 testing kits are readily accessible in health care facilities. Thus it is imperative to define a fully automated detection method as an instant alternate treatment possibility to limit the occurrence of COVID-19 among individuals. In this paper, a two-step Deep learning (DL) architecture has been proposed for COVID-19 diagnosis using CXR. The proposed DL architecture consists of two stages, "feature extraction and classification". The "Multi-Objective Grasshopper Optimization Algorithm (MOGOA)" is presented to optimize the DL network layers; hence, these networks have named as "Multi-COVID-Net". This model classifies the Non-COVID-19, COVID-19, and pneumonia patient images automatically. The Multi-COVID-Net has been tested by utilizing the publicly available datasets, and this model provides the best performance results than other state-of-the-art methods.

1. Introduction

Coronavirus diseases (COVID-19) outbreak the health of the population of the world due to the infection of patients with severe acute respiratory syndrome corona virus2 (SARSCoV2) [1]. The virus causes respiratory infection with signs including cough, fever, and breathing difficulty in many more extreme cases. Although the virus first began in Wuhan, China, in November 2019, the virus is spreading far and wide so rapidly, with 0.739 million deaths and more than 20 million individuals around 188 nations in the world as of October 2020 [2]. It is one of the fastest progressive diseases ever seen. Since it is a new viral infection and shifts accumulation rapidly, there is no practice standards for the evaluation or examination procedure and appropriate treatment. Thus, COVID-19 may infect a very large group of people unless prevention practices are implemented. The Reverse transcription-polymerase Chain Reaction (RT-PCR) is the common testing procedures of COVID-19 infected patients [3]. The time required to confirm COVID-19 infected patients is high using these tests, and it has given a number of false-positive results [4]. Hence there is an urgent need for an automated COVID-19 diagnosis system.

The medical imaging modalities such as chest Computed Tomography (CT) and Chest X- Ray (CXR) images are used to diagnose the COVID-19. The CXR are utilized in many hospitals because X-imaging modality is very simple and easily available. In recent years Deep Learning (DL) produced better results in radiological based medical imaging diagnosis. Among DL algorithms, Convolutional Neural Network (CNN) is widely used in many medical image classifications [5]. These are the motivational factors to propose this paper.

In medical image classification, the optimization algorithm plays a vital role in minimizing the cost function and maximizing efficiency [6]. In CNN training, hyperparameters selection decides the speed and maximize the performance of the training process [7]. Many COVID-19 diagnosis works are reported manual hyperparameters selection, which leads to more training time [8–10]. To address these issues, this paper proposes an optimized hyperparameters selection of the CNN model using the "Multi-Objective Grasshopper Optimization Algorithm (MOGOA)". Other

* Corresponding author.
E-mail address: murugan.rmn@ece.nits.ac.in (R. Murugan).

https://doi.org/10.1016/j.asoc.2021.108250
1568-4946/

optimization algorithms limitation is their dynamicity, uncertainty, and work for only a single objective.

This paper proposes an ensemble network that includes InceptionV3 and ResNet-50 architecture to diagnose COVID-19 from the CXR images automatically. This ensemble network hyperparameters are optimized using the Multi-Objective Grasshopper Optimization Algorithm (MOGOA); hence this system is named "Muti-COVID-Net". This system automatically extracts the features and classifies them from the COVID-19, Non-COVID-19, and pneumonia CXR images. The superiority of the proposed Muti-COVID-Net is to produce an efficient COVID-19 diagnosis system and built the Optimum performance through experimentation.

The novelty of the work is the proposed deep ensemble learning architecture whose hyperparameters are optimized using MOGOA optimization to accurately diagnose COVID-19 from CXR images. Advantage of using ensemble model is to reduce the network's variance for the fast random generation of weight and biases of the networks. Also, the proposed ensemble deep neural networks provide better classification performance and generalize the results than single DNN. MOGOA optimization will tune the hyperparameters of the network to give best accuracy, sensitivity and specificity.

The contributions of this work as follows:

I. To overcome the RT-PCR false-negative results, CXR images are utilized to diagnose COVID-19.
II. This paper proposes an ensemble model, consisting of InceptionV3 and ResNet50 pre-trained networks for COVID-19 diagnosis.
III. Performance of ensemble deep learning network highly depends on the selection of the hyperparameters. Improper selection of hyperparameters of the ensemble network may lead to more false-negative. Hence MOGOA has been used in this paper to optimize the hyperparameters ensemble network.
IV. Extensive performance and comparative analysis has been made in this paper to test the performance of the Multi-COVID-Net with the help of various performance metrics.

This paper's remaining sections are arranged as follows: The detailed literature survey has presented in Section 2. Section 3 presents the need, motivation, and theoretical, mathematical fundamentals of MOGOA and the proposed method's workflow. The experiments, including results, discussion, performance, and comparative analysis, have been described in Section 4. Finally, Section 5 presents the concluding remarks of this paper.

2. Related works

In recent times DL algorithms have been used for medical image classification. Such algorithms have been also used for classifying COVID-19 in Chest X-ray images. This section presents a detailed literature review of DL algorithms for diagnosing COVID-19.

Hassantabar et al. [5] presented two algorithms for COVID-19 diagnosis, including fractal-based feature extraction on DNN and CNN [5]. CNN achieved better accuracy (93.2%) then DNN (83.4%). However, this study has used only 682 images for both training and testing. Altan and Karsan [11] have proposed a hybrid model that includes a 2D curvelet transform, Chaotic slap Swarm Algorithm (CSSA), and DL. The curvelet transform applied to obtain the feature matric from the X-ray images; then, the feature matrix has optimized using CSSA, and EfficientNet-B0 DL architecture used for the classification of COVID-19 images. However, this method used 159 images for their experiments. The authors of [12] proposed a Deep Transfer Learning (DTL) method for automatic COVID-19 diagnosis. They used the advanced version of Inception called Xception to train and test the networks. This works showed promising results, but it has a lack of clinical explanation.

Imran et al. [13] attempted to diagnose the COVID-19 from cough sounds. The coughing sound was recorded by a smartphone application and test by their developed model. Promising results were achieved but with the help of cough sound, identifying COVID-19 is quite difficult. Since cough is the general symptoms for other than COVID-19. The DNN based architecture called "CoroNet" was developed by Khan et al. [14] for classifying four-class cases that include normal, COVID-19, Pneumonia, and Pneumonia bacterial. They have used the Xception network for training and testing the images. However, this work has not considered viral pneumonia images.

A deep CNN-based network called "ConXNet" was proposed by Mahmud et al. [15] This network has analyzed the "depth wise convolution" for extracting the feature from X-ray images. This study also designed a different form of "ConXNet" model to test different X-ray images' resolutions. However, the processing time and computational complexity were high. Minaee et al. [16] developed two-class classification method such as Non-COVID-19 and COVID-19 using deep TL used four different CNN models, including Sequeezenet, DenseNet-121, ResNet18, and ResNet50 pre-trained architectures. The SqueezNet achieved good results of sensitivity and specificity among other networks. Nour et al. [17] proposed a novel CNN model and trained from scratch instead of using TL. They designed five-layer architecture to extract the features from CXR images then the features were fed into ML algorithms then the features were optimized by the Bayesian algorithm.

The CNN-based "DarkCovidNet" was proposed by Ozturk et al. [18] from the CNN algorithm. They developed two classification architectures: binary and multi-class. The binary includes Non-COVID-19 and Non-COVID-19, and the multi-class, including COVID-19, Non-COVID-19, and pneumonia. The authors demonstrated that the binary architecture achieved better results than multi-class architecture. However, this method suffers from the lack of clinical intervention. The authors in [19] employed the deep TL for fast and accurate diagnosis of COVID-19. This study has been used to differentiate the COVID-19 and COVID-19 images and this method has been produced an accuracy of 95.61%. Vaid et al. [20] proposed CNN architectures to improve the best accuracy of COVID-19 confirmed cases from CXR imaging. This has a lack of implementation issues.

Narin et al. [21] utilized three CNN, which includes InceptionV2, InceptionV3, and ResNet50, with 100 CXR images out of which 50 are normal and 50 are COVID-19. ResNet50 provided an accuracy as high as 98%. This work's limitation is the limited number of data, which are only 100 X-ray images collected from the Github repository. Deep feature extraction was based on eleven pre-trained models by Sethy et al. [4]. In this work, the accuracy obtained in ResNet50 is high compared to other pre-trained models. The limitation of this work is a lack of data availability and high computational cost.

Hemdan et al. [22] proposed an ensemble architecture called "COVIDX-Net", which incorporates the pre-trained models, such as the modified VGG19 and the Google MobileNet. Each DNN model can analyze the relatively stable levels of intensity of the CXR image to categorize the patient's condition as either COVID-19 or Non-COVID-19. Apostolopoulos et al. [23] trained the CNN models to get more accuracy. The CNN architecture has been produced 93.48 accuracy for classifying COVID-19 and normal CXR images. The major limitations of this work are lack of data availability and high computational cost.

Table 1
State-of-the-art DL model for COVID-19 diagnosis system through CXR images.

Author	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F-score (%)
Hassantbar et al. [5]	93.2	96.1	–	–	–
Altan et al. [11]	95324	93.61	96.05	92.22	92.91
Das et al. [12]	97.4	97	97.2	–	96.9
Imran et al. [13]	88.76	91.71	95.27	86.60	89.08
Khan et al. [14]	89.5	97	100	–	–
Mahmud et al [15]	97.4	–	–	–	–
Minalee et al. [16]	–	98	–	–	–
Nour et al. [17]	98.97	89.37	99.75	96.72	–
Oztunk et al. [18]	87.02	–	–	–	–
Panwar et al. [19]	95.61	–	–	–	–
Vaid et al. [20]	96.3	–	–	–	–
Nain et al. [21]	98	–	–	–	–
Sethy et al. [4]	95.38	–	–	–	–
Hemdan et al. [22]	90	–	–	–	–
Wang et al. [10]	83.5	–	–	–	–
Alazab et al. [3]	94.80	–	–	–	–
Proposed	98.21	99.63	97.59	95.39	97.46

Wang [10] introduced "COVID-Net", a Deep CNN model for COVID-19 screening from CXR images. This model also evaluated how "COVID-Net" experimentally verified using the data annotation strategy in an effort not only to achieve a useful understanding of the key factors related to COVID-19 cases, which can help clinicians to enhance evaluation. The need for CXR images was strongly advised by Alazab et al. [3] to screen COVID-19 since X-rays machines are readily available in nearby healthcare centers reasonably quickly and at a minimal price. They proposed the VGG19 CNN model and obtained promising results than other methods.

The summary of these related works is presented in Table 1. The transfer learning techniques are implemented [3–8] by many researchers and data scientists based on predefined CNN architectures such as ResNet50, InceptionV3, Inception-ResNetV2, VGG-19, Mobile-Net, Xception to obtain better performance. However, such CNN models require more time for training and testing. The gap is that the methods proposed did not discuss parameter tuning, which has the potential to substantially impact the accuracy. The COVID-19 patient database used is small in the literature reviewed. The number of images, both training, and testing, may be increased to get better performance metrics. However, as these metrics extensively depends on the number of samples representing each category of images. More COVID-19 patient data and more computational resources are required, then the only generic model may be developed.

3. Methodology

The proposed model performed the automatic classification of CXR images into Non-COVID-19, COVID-19, and pneumonia patients. The proposed architecture is based on the ensemble of deep learning networks to reduce the variance because of the networks random generation of weight and biases.

The ensemble of deep neural networks (DNN) also provides better performance and generalization results than a single DNN. The performance of DNN networks greatly depends on the selection of hyperparameters. Therefore, both DNN networks' hyperparameters are optimized using MOGOA, optimizing network hyperparameters for maximum accuracy, sensitivity, and specificity. Two DNN networks used for COVID-19 diagnosis are Inceptionv3 and ResNet50. First, both networks' hyperparameters are optimized for COVID-19, normal, and pneumonia dataset, and then features are extracted from both networks. These features are combined and fed to "support vector machine (SVM)" classifier for classification into the appropriate label. Fig. 1 shows the flow diagram of the proposed model for the diagnosis of COVID-19.

3.1. Ensemble learning

Deep learning networks are nonlinear that use a large amount of training data and tune the weights and parameters using a stochastic learning and analytical approaches. Because of the stochastic training method, every iteration will give different sets of weights and biases, which results in different prediction results. This is referred to as variances, the fluctuations in the performance results due to change in the parameters of DNN. Ensemble learning is used to resolve the high variance, which trains the many neural networks (NN) parallel to solve the classification task. Ensemble learning is first proposed by Hansen and Salamon [24] to train a finite number of NN for the same problem. Ensemble learning not only improves the variance of the network but also improves the generalization capability of NN.

In the proposed methodology, Inceptionv3 and ResNet50 DNN are used for ensemble networks. Both networks are trained parallel on the CXR image datasets for COVID-19, Non-COVID-19, and pneumonia. After training, features of both networks are combined and classified using the SVM classifier.

3.2. Multi Objective Grasshopper Optimization Algorithm (MOGOA)

3.2.1. Need of optimization

To tune the weights and biases of ensemble DNN, hyperparameters values such as batch size, learning rate, momentum, number of epochs, and regularization coefficient should be chosen very carefully. These values influence the gradient descent training, so it is important to optimize them. Improper selection of the hyperparameters may lead to poor results, while changing the hyperparameters and checking the results will take much time for ensemble DNN. Therefore, there is a need to optimize hyperparameters, which automatically find optimal values and keeps updating them according to the training data. The optimization goal is to train the DNN for the best accuracy, sensitivity, and specificity.

3.2.2. Motivation of using MOGOA

In the field of optimization, nature-inspired algorithms become very popular for solving different challenging problems. The most popular nature-inspired optimization is the Genetic Algorithm (GA) inspired by Darwin's theory of evolution [25]. GA creates a set of random solution and then search for the best solution using mutation and crossover. The main drawback of GA optimization is its binary nature and difficulty of problem formulation for this algorithm. The next group of nature inspired optimization includes "Particle swarm optimization (PSO) [26]", "Whale Optimization algorithm (WOA) [27]", etc.

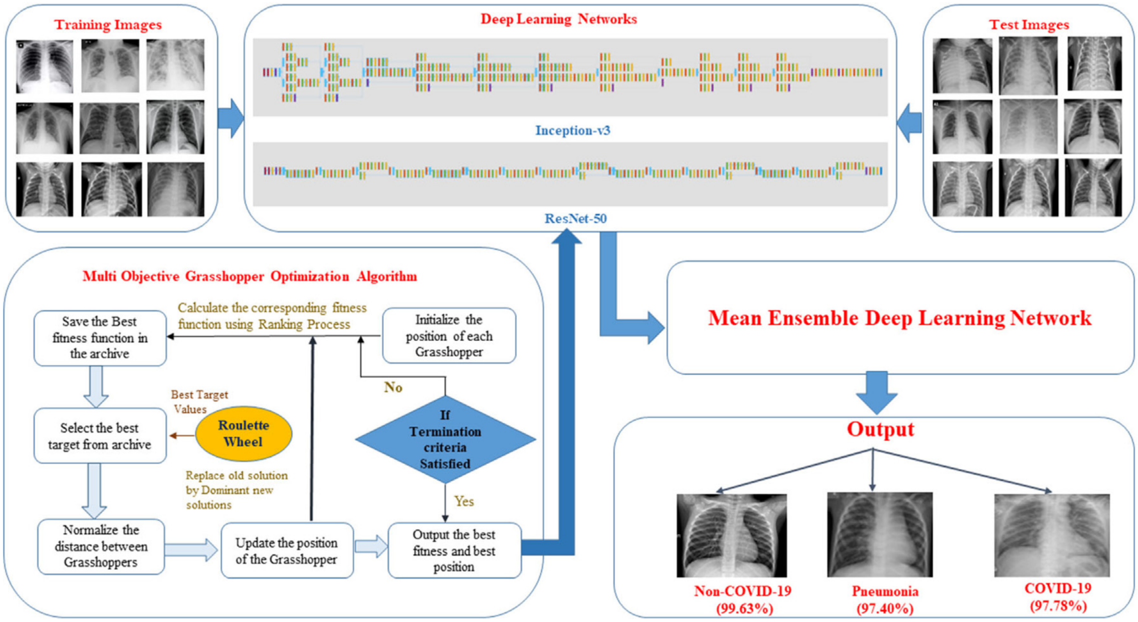


Fig. 1. Workflow of Multi-COVID-Net architecture.

The optimization process using such optimization algorithms is divided into two phases: exploration and exploitation. Exploration refers to the algorithm's capability to have more diverse values, while exploitation refers to searching in small spaces. The "Grasshopper optimization algorithm (GOA) [28]" mimics the behavior of the swarm of grasshoppers (GH). GOA is able to balance between exploration and exploitation. GH first move locally to search for prey and then move freely to explore in diverse space. However, in the previously mentioned optimization algorithms, exploration is done previous to exploitation. They first find the regions in the large feature space (exploration) and then search locally (exploitation) to find the objective function's approximate solution. GOA has been extended to a multi-objective version called Multi-Objective GOA (MOGOA). It has been demonstrated that MOGOA can be very efficiently and superior of a wide range of problems compared to other existing algorithms. This motivated our attempts to employ MOGOA for optimizing the selection of hyperparameters of ensemble deep learning networks.

3.2.3. Mathematical background of MOGOA

GOA is motivated by the nature-inspired swarming nature of GH [29]. The GH position in the GOA algorithm denoted the possible solution and represented using Eq. (1).

$$P_i = S_i + G_i + W_i \tag{1}$$

Where, S_i denotes the social interaction, G_i denotes the gravity force and W_i denotes the wind advection of the ith GH in the swarm. For optimization problem, mathematical model proposed for GOA [29] is given in Eq. (2).

$$P_i^d = c\left(\sum_{\substack{j=1\\ j\neq i}}^{N} c\,\frac{u_{bd}-l_{bd}}{2}\,s\left(\left|x_j^d - x_i^d\right|\right)\frac{x_j - x_i}{d_{ij}}\right) + \widehat{T_d} \tag{2}$$

Where, u_{bd} denotes the upper bound and l_{bd} denotes the lower bound of dth dimension. T_d represents the target and c is a coefficient of decreasing order to shrink the area of comfort, repulsion and attraction. In Eq. (2), c is the very important parameter that should be updated for each iteration to make the balance between exploration and exploitation.

In the proposed work, accuracy, sensitivity, and specificity of the COVID-19 diagnosis capability of ensemble DNN are the objective to optimize the hyperparameters of DNN. Eq. (3) shows the formulation of the proposed multi-objective problem.

$$\begin{aligned}
&\text{Minimize}: f_m(x), \quad m = 1, 2, 3\\
&\text{Subject to}: g_i(x) \geq 0; \quad i = 1, 2, \ldots, k\\
&\qquad h_j(x) = 0; \quad j = 1, 2, \ldots, l\\
&\qquad x_i^{\mathrm{lb}} \leq x_i \leq x_i^{\mathrm{ub}}, \quad i = 1, 2, \ldots, n
\end{aligned} \tag{3}$$

Where, m is the total number of objectives, n is the number of variables to be optimized, k is the number of inequality constraints and l is the number of equality constraint. g_i indicates the ith inequality constraint, h_j is the jth equality constraint and $[x_i^{\mathrm{lb}}, x_i^{\mathrm{ub}}]$ indicate the boundaries of ith variable.

MOGOA generate more than one solution to the optimization problem. Therefore, multiple solutions are compared using a Pareto Optimality (PO) operator. In MOGOA, first, the best PO solution is stored in an archive. Next, the selection of the target is the main challenge in multi-objective optimization. In MOGOA, the target is selected from the set of PO operators. For selecting a target, the number of neighborhood solutions of each PO operator stored in the archive is counted, used as the quantitative measure. The probability of selecting the target based on the number of neighborhood solutions is Eq. (4).

$$p_i = \frac{1}{N_i} \tag{4}$$

N_i denotes the number of solutions for the ith PO operator. Then Roulette wheel is used to select the target using Eq. (4). To update the archive regularly, if external solution is dominated then solution in the archive is replaced by an external solution. The Pseudo code of the proposed MOGOA based Multi-COVID-Net is given in Algorithm (1).

In the following section, the experimental results are presented, discussed, and analyzed.

4. Experiments

This section presents the experiments conducted to prove the efficiency of the proposed Multi-COVID-Net based on different classifications and optimization algorithms.

Algorithm 1 Pseudocode of the MOGOA Optimized CNN

Require: *Training Set T; Testing Set Tt*; Number of classes C; *hyper parameters*: $K1, K2, K3, K4, K5, K6$;
hyper parameter evaluation function: L_D

Initialize: *Random Grasshopper Population* X_i $(i = 1,2, \dots n)$, *dimensions*: d; *number of optimization iterations*: t;
Objective functions: Accuracy, Sensitivity, Specificity

Procedure: for training data T
Initialize: Inception-v3 DL Network
Resnet50 DL Network
for *number of optimization iterations*, t

Calculate fitness function
Select the best target using Roulette wheel from the archive
Compare if fitness value is better than the best target value
Yes: Update the values of all Grasshoppers using $P_i^d = c\left(\sum_{\substack{j=1 \\ j \neq i}}^{N} c \frac{u_{bd} - l_{bd}}{2} s\left(\left|x_j^d - x_i^d\right|\right) \frac{x_j - x_i}{d_{ij}}\right) + \widehat{T_d}$

No: Keep the previous values

If the archive is full
Delete some solutions to hold new solutions.
Applying Roulette wheel and use $p_i = \frac{1}{N_i}$
end If
If any new added solutions to the archive are outside boundaries
Update the boundaries to cover the new solution(s).
end If

$(K1, K2, K3, K4, K5, K6) = hyper\ parameter\ evaluation\ function(L_D)$

end
$w \,\&\, b \leftarrow Sgdm(L_D, K1, K2, K3, K4, K5, K6)$
Train DL Networks partially using optimized Parameters
Ensemble features of both DL Networks
Test on Training data and calculate Accuracy, Sensitivity, Specificity

end Procedure

Table 2
Details of dataset.

Dataset	Images category	No. of images
Chest imaging [30]	COVID-19	134
SIRM COVID-19 database [31]	COVID-19	64
COVID-chest X- ray [32]	COVID-19	646
Fig. 1 COVID-19 Chest X-ray [33]	COVID-19	55
Provincial peoples hospital [34]	COVID-19	1
Kaggle [35]	Normal	900
Kaggle [35]	Pneumonia	900
	Total	2700

4.1. Dataset

The publicly available CXR images of Non-COVID-19, COVID-19, and pneumonia are used in this work. These investigations have been driven all together of 2700 images, in which there 900 images for each class. The description of the dataset is given in Table 2.

4.2. Hardware and software details

The proposed architecture has been implemented in MATLAB R2020b (version 9.7) and executed using system configuration of NVidia, Windows 10 Pro, and 64 GB RAM GPU.

4.2.1. Training

The proposed Multi-COVID-Net architecture has been trained by 70% of randomly partitioned images, and these images are resized into 224 X 224 X 3 for better resolution. The proposed architecture classifies the image in either of the three categories: COVID-19, Non-COVID-19, and Pneumonia. The InceptionV3 and ResNet50 pre-trained CNN architecture are used to diagnose these images. These networks hyperparameters are optimized using MOGOA, and the parameters are presented in Table 3. Fig. 2 shows the sample training images of Multi-COVID-Net.

Table 3
Training options using MOGOA optimization.

Parameters	InceptionV3 Rate	ResNet50
Momentum	0.5224	0.750
Learning rate	0.0549	0.0650
Epoch	7	6
L2Regularization	1.3042e−04	1.1746e−04
Batch size	32	32

4.2.2. Testing

The proposed Multi-COVID-Net architecture has been tested by randomly selected 30% dataset images. These images are resized into 224 X 224 X 3 before testing of the network. After that, these image are fed into a trained Multi-COVID-Net model. The network architecture firstly extracts the features and classify the images using fully connected and soft-max layers. Fig. 3 shows the sample testing images.

4.3. Performance parameters and evaluation metrics

In this work, accuracy, sensitivity, specificity, precision, and F1 score values are utilized to determine the classification results. The definition of these parameters are represented in Eqs. (5), (6), (7), (8), and (9) [36] and the related terms are presented Table 4.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{5}$$

$$\text{Sensitivity or recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{7}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{8}$$

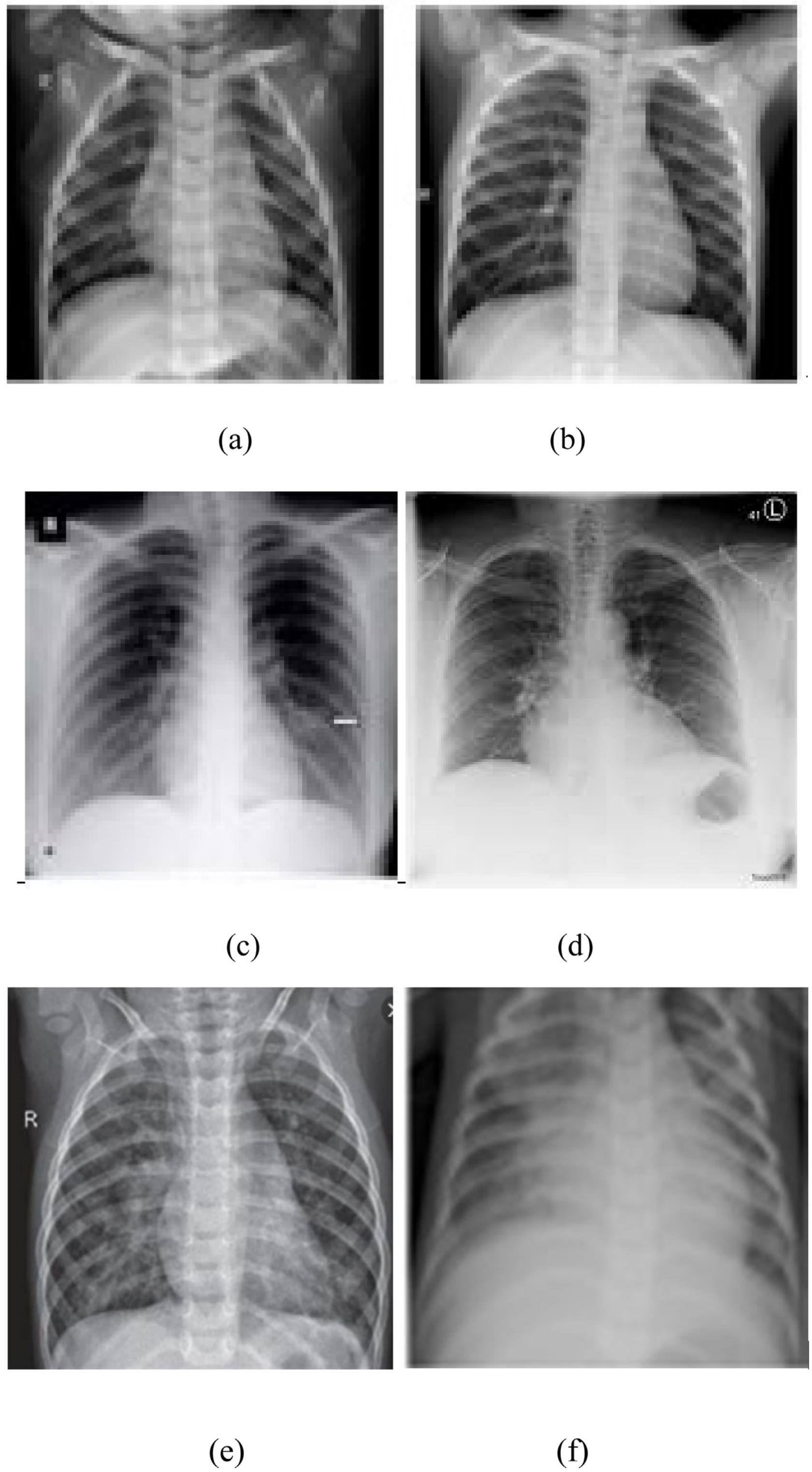


Fig. 2. Training image samples (a–b) Non-COVID-19 (c–d) COVID-19 (e–f) Pneumonia.

Table 4
Performance parameters.

Parameters	Definition
True Positive (TP)	The COVID-19 image is correctly diagnosed
False Positive (FP)	The Non-COVID-19 image is misdiagnosed as COVID-19
True Negative (TN)	The Non-COVID-19 image is correctly diagnosed
False negative (FN)	The COVID-19 image is misdiagnosed as non-COVID-19

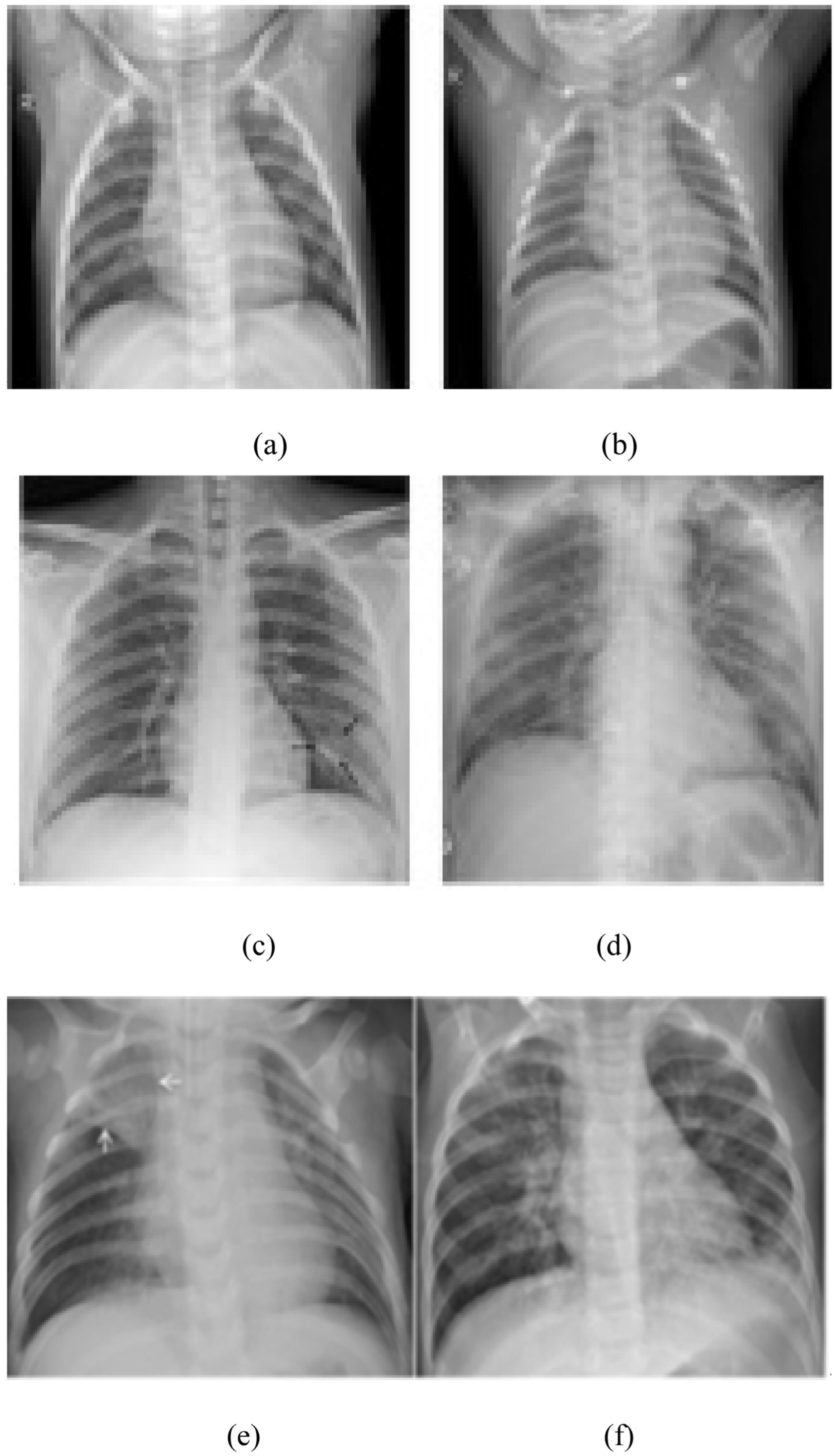


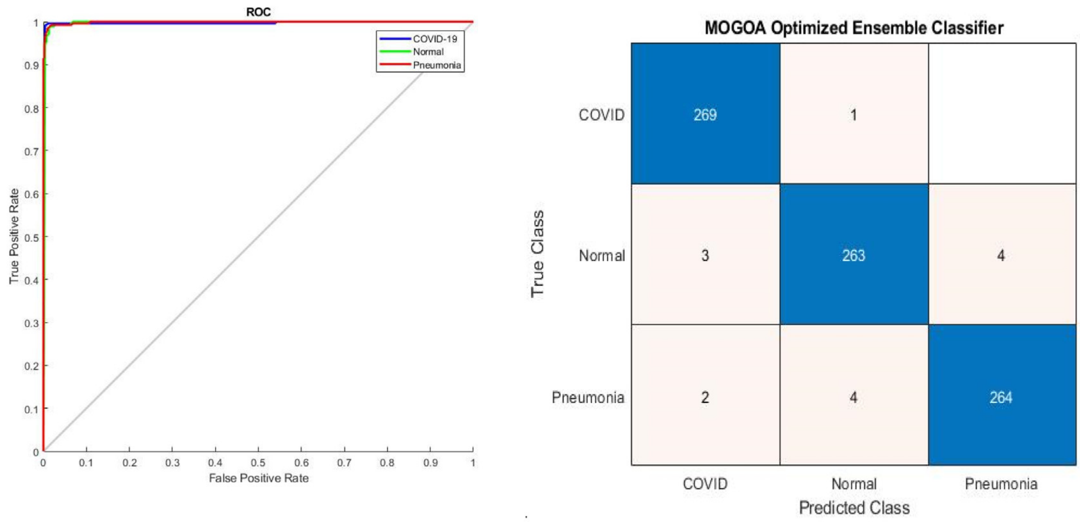
Fig. 3. Testing image samples (a–b) Non-COVID-19 (c–d) COVID-19 (e–f) Pneumonia.

$$\text{F1 Score} = 2\frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{9}$$

In addition to the above performance metrics, Receiver Operating Characteristics (ROC) and Confusion Matrix (CM) are also used to evaluate the proposed work. The ROC is the representation of the characteristics of classification methods performance execution. This ROC has generated between specificity vs. sensitivity. CM is describing the performance of classification in terms of tabular form.

4.4. Experimental results

In this sub-section, the results and the performance of the proposed method is presented. The trained Multi-COVID-Net model



(a) (b)

Fig. 4. (a) Generated ROC (b) Generated CM.

Table 5
Results of the Multi-COVID-Net model.

Accuracy	Sensitivity	Specificity	Precision	F1 Score
98.27%	99.63%	97.59%	95.39%	97.46%

Table 6
Parameters of MOGOA.

Parameters	Values
Dimension	8
Objective function	3
Number of iterations	20
Population size	200
Archive size	32
Lower bound	[0.2, 0.01, 1.0000e−04, 0.2, 0.01, 1.0000e−04, 8]
Upper bound	[0.9, 0.1, 2.0000e−04, 0.9, 0.1, 2.0000e−04, 64]

results are presented in Table 5, showing better performance values.

4.4.1. ROC and CM

The ROC and CM have been generated to test the performance of the Multi-COVID-Net model classification effectiveness. The proposed Multi-COVID-Net produces better accuracy in COVID-19 (98.88), Non-COVID-19 (97.40), and Pneumonia (97.77) classes. Hence this Multi-COVID-Net may be utilized for automatic screening of COVID-19. The generated ROC and CM are shown in Fig. 4. The selected parameters of the MOGOA is presented in Table 6.

4.4.2. Performance analysis with other classifiers

In this work, six standard DL networks are used for the classification of the images. These are Decision Tree (DT) [37], K-Nearest Neighbor [38], Support Vector Machine (SVM) [39], Navies Bayes (NB) [40], Random Forest (RF) [41], CNN [42], Stack Encoder (SE) [43], ResNet50 [44], InceptionV3 [45], Ensemble Deep Learning Network (EDLN) [24] Without Optimization and proposed Ensemble Deep Learning Network (EDLN) with MOGOA Optimization algorithms. The performance comparison of these networks has done using metrics: accuracy, specificity, sensitivity, precision, and F1 score, is given in Table 7. The performance comparison in terms of the ROC and CM is illustrated in Figs. 5 and 6. ROC curve is the graphical representation to show the classification ability of the network. The more the curve closer to top left border of ROC space, results will be more accurate. Fig. 5 compares the ROC curve of the proposed classifier with the state of the art classifiers and shows the closest ROC curve toward the top left of the ROC curve to prove its efficiency. From all these comparisons, the optimized MOGOA model gives the best accuracy compared to other DL networks.

Table 7
Performance comparison of classifiers.

Method	Accuracy	Sensitivity	Specificity	Precision	F1 Score
DT	71.98	77.04	69.44	55.76	64.70
K-NN	73.95	96.67	62.59	56.37	71.21
SVM	77.16	94.07	68.70	60.05	73.30
NB	85.06	87.78	83.70	72.92	79.66
RF	86.67	84.41	86.30	76.13	81.38
CNN	91.96	87.04	95.93	91.44	89.18
SE	75.19	95.59	66.48	58.00	71.33
ResNet50	88.27	97.78	83.52	74.79	84.75
InceptionV3	91.85	93.33	91.11	84.00	88.42
EDLN	94.19	97.78	93.89	88.89	93.12
Proposed	98.27	99.63	97.59	95.39	97.46

4.4.3. Performance analysis with other optimization algorithms

This sub-section presents the performance comparison of different optimization algorithms with MOGOA. The optimization algorithm are GA [46], Pattern Search (PS) [47], PSO [26], WOA [27], Grasshopper Optimization Algorithm (GOA) [48], Multi-objective Genetic Algorithm (MOGA) [28]and EDLN based Multi-objective GOA (MOGOA). Table 8 shows the performance comparison of optimized networks. From all these comparisons, MOGOA optimized Multi-COVID-Net provides the best performance metrics in terms of accuracy, sensitivity, specificity, Precision and F1 score compared to other optimization techniques. Therefore, we offer the proposed method as an accurate tool for screening COVID-19 patients.

4.4.4. Performance analysis with grid search optimization

Grid search strategy (GSS) is the conventional and most popular method used for optimization of the hyperparameters of DL networks. GSS can be used to optimize only single objective

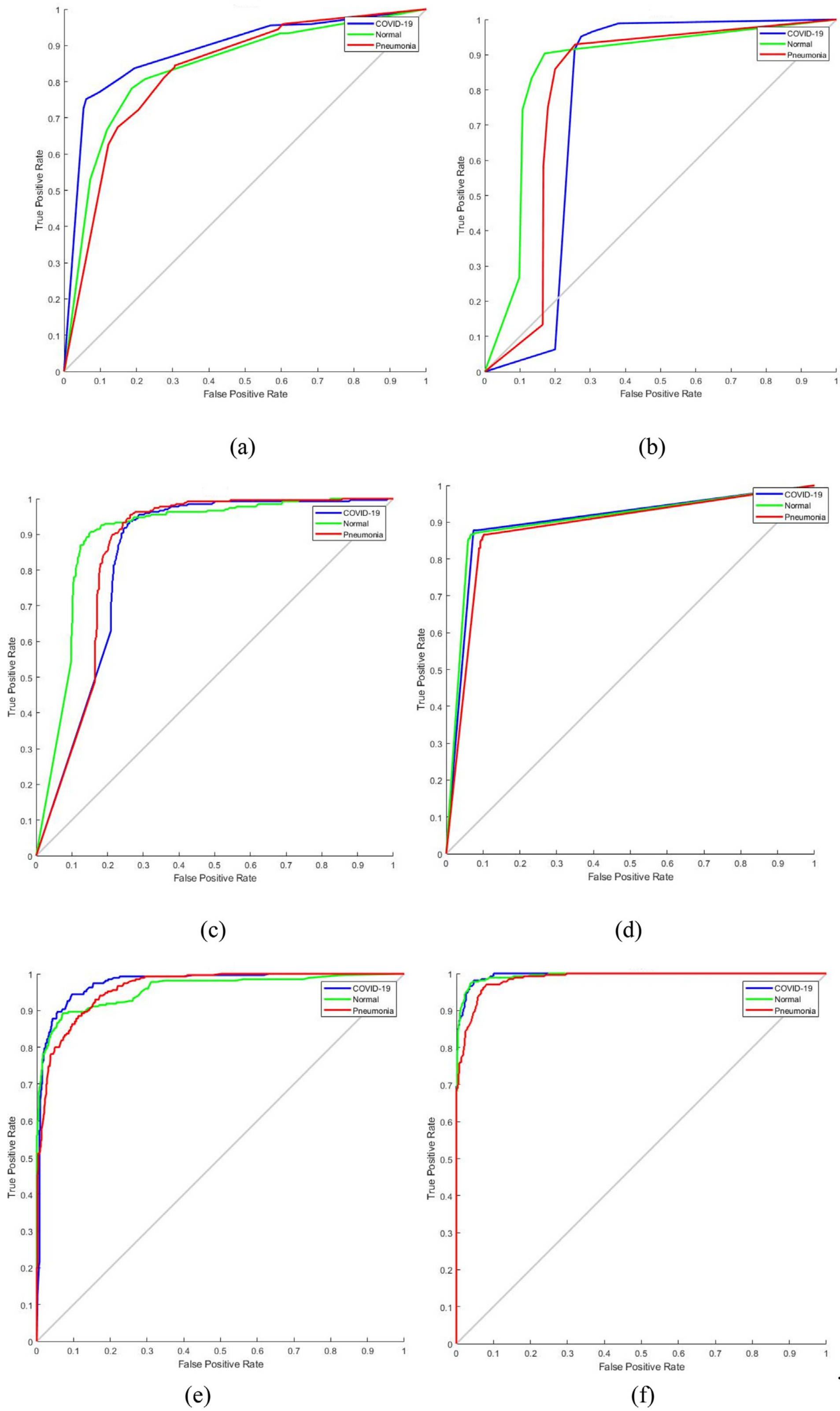
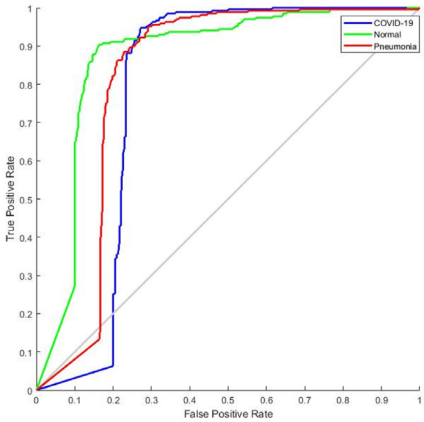
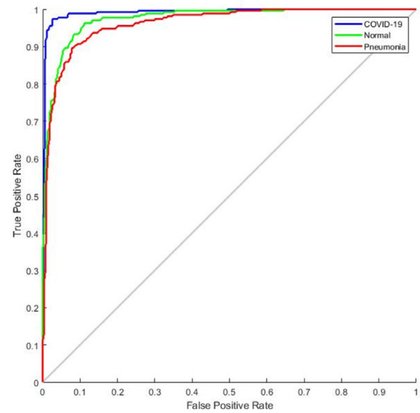


Fig. 5. ROC of (a) DT (b) K-NN (c) SVM) (d) NB (e) RF (f) CNN (g) SE (h) ResNet50 (i) InceptionV3, (j) EDLN Without Optimization (k) EDLN with MOGOA Optimization algorithms.

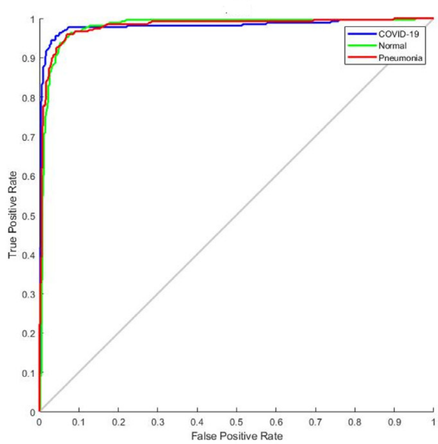
and hyper parameter space is specified manually to optimize the single objective. GSS creates the grid of all the possible solutions of specified hyperparameters space. The combination which gives the best results from the grid, will be selected hyperparameters.



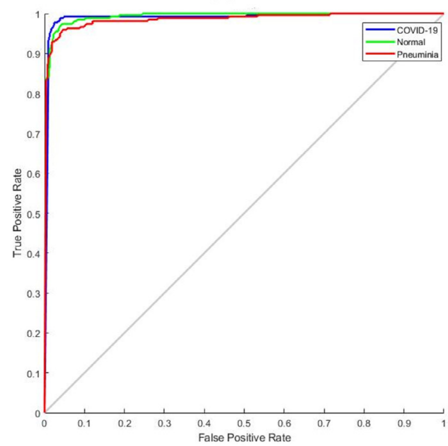
(g)



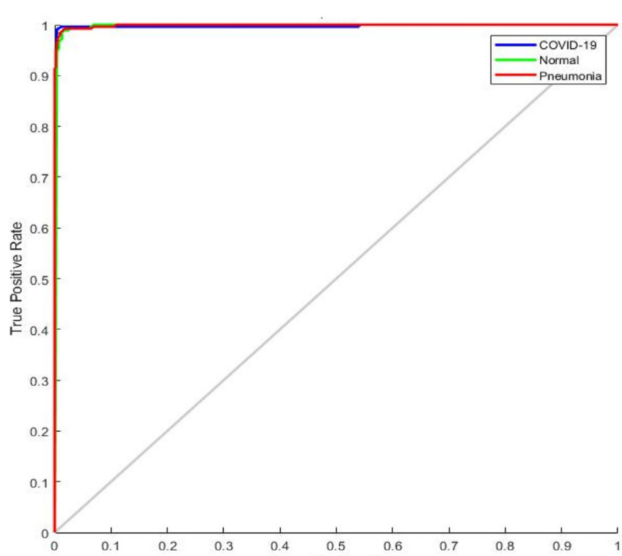
(h)



(i)



(j)



(k)

Fig. 5. (*continued*).

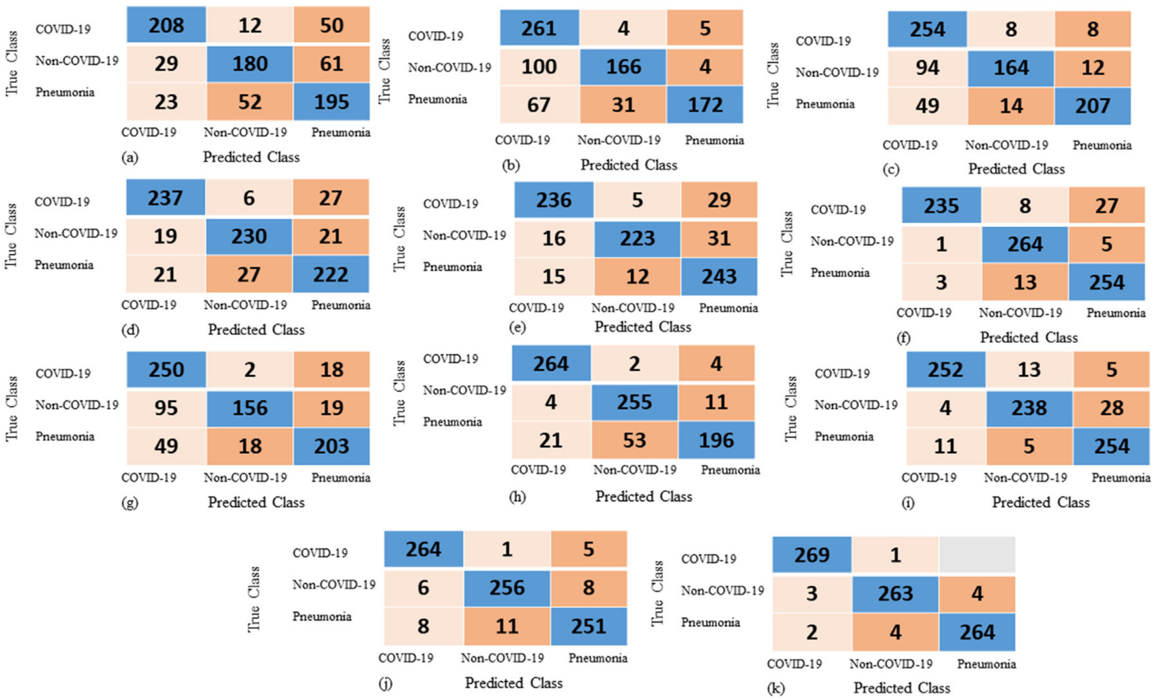


Fig. 6. Confusion Matrix of (a) DT (b) K-NN (c) SVM) (d) NB (e) RF (f) CNN (g) SE (h) ResNet50 (i) InceptionV3, (j) EDLN Without Optimization (k) EDLN with MOGOA Optimization algorithms.

Table 8
Performance comparison of optimization algorithms.

Method	Accuracy	Sensitivity	Specificity	Precision	F1 Score
GA	95.80	97.04	95.19	90.97	93.91
PS	94.95	95.56	93.15	87.46	91.33
PSO	95.19	96.30	94.63	89.97	93.02
WOA	94.44	97.78	92.78	87.13	92.15
GOA	96.30	98.15	95.37	91.38	94.64
MOGA	96.53	98.63	96.48	93.40	96.42
Proposed	98.21	99.63	97.59	95.39	97.46

Although GSS strategy is simple but the main drawback is the increase in the number of iterations exponentially with the addition of each parameter. Comparison of GSS with non-optimized and MOGOA optimized network is sown in Fig. 7 with the variations in the percentage of training images.

4.4.5. Performance analysis with another dataset

To observe the scalability and generalization of the proposed architecture, another dataset is used to check the performance. Another dataset is taken from the web link https://github.com/armiro/COVID-CXNet and https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia which includes 200 images of each class: COVID-19, pneumonia, and normal. The test dataset is directly given as input to the proposed model. Model first do the pre-processing and resizing of the input image to make it fit to the deep ensemble model. Hyper parameters of the DL networks will be fixed which are optimized using MOGOA. Table 9 shows the performance comparison of the optimized and non-optimized ensemble DL network in terms of accuracy, specificity, sensitivity, precision and F1-score.

4.5. Comparative analysis using cross-validation

Cross-validation (CV) is a re-sampling technique; it helps to avoid over-fitting or under-fitting. It is utilized to evaluate the proposed model with small data samples. The X-ray images have split into training and testing samples and referred to the parameter k times. As such, this procedure is called k-fold cross-validation. In this method, the k value is chosen as 10 (into 10 iterations). The experiments have been conducted between optimized and non-optimized networks using the calculated error rate k vs. a number of iterations. The performance of this 10-fold cross-validation network is presented in Fig. 8. From these experiments, the proposed optimized network produces a very less error rate than the non-optimized network.

4.6. Computational complexity

The computational complexity of the ensemble network depends on the number of deep learning architectures used for the classification. The complexity of ensemble network is calculated using $O(2^N)$, where N is the total number of the networks. In the proposed algorithm, two DL networks are ensemble, therefore the value of $N = 2$. Computational complexity of MOGOA optimization is $O(MP^2)$, where M is the number of objectives and P is the number of solutions. Therefore, total computational complexity of the proposed network is $(O(2^N) + O(MP^2))$.

4.7. Discussion

The millions of individuals are infected by COVID-19; hence there is an urgent need for accurate screening of COVID-19. This work has been diagnosed with COVID-19 patients from non-COVID-19 and pneumonia patients through CXR images. CXR is an affordable and commonly available medical imaging modality in all radiological centers and hospitals. The developed Multi-COVID-Net was trained by 1890 numbers of CXR images, providing the best performance metrics.

The Multi-COVID-Net model performance was compared with other state-of-the-art DL networks. Also, the optimized Multi-COVID-Net model has been compared with metaheuristic optimization algorithms to prove its efficiency. The proposed Multi-COVID-Net architecture has the ability to diagnose COVID-19 patients without trained radiologists. The advantages of

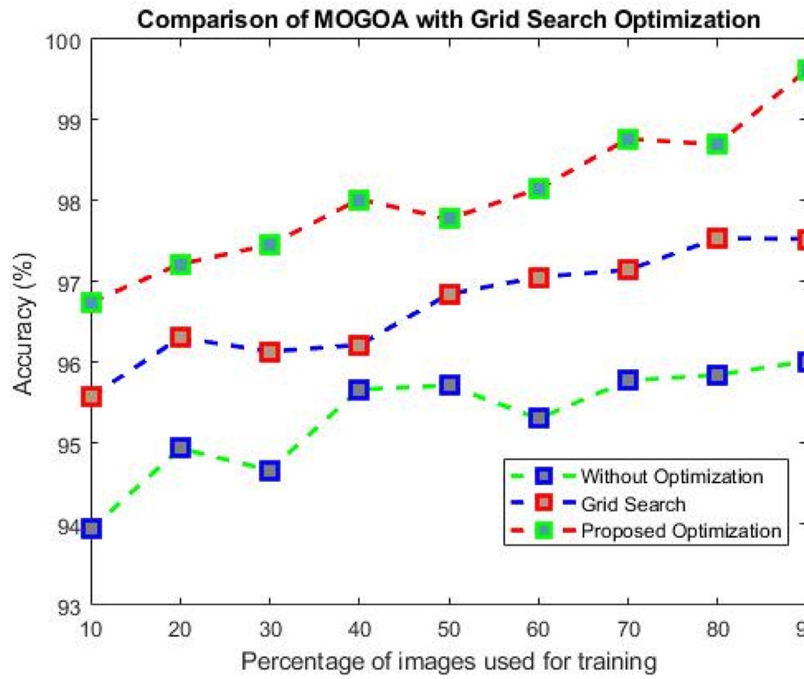


Fig. 7. Comparison of MOGOA with GSS.

Table 9
Performance comparison on another dataset.

Method	Accuracy	Sensitivity	Specificity	Precision	F1 Score
Without optimization	91.00	99.50	86.75	78.97	92.91
With Proposed optimization	95.33	99.50	93.25	88.05	93.43

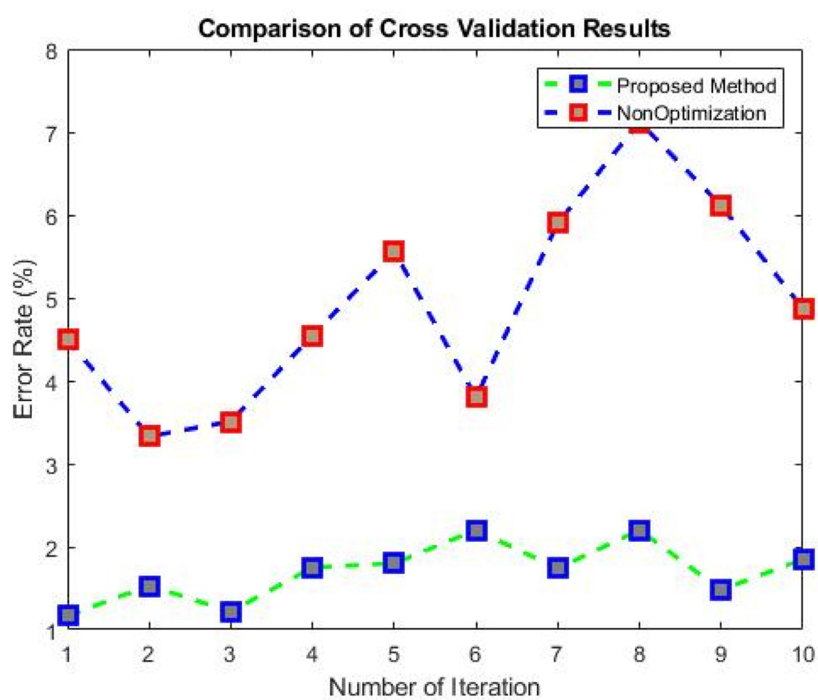


Fig. 8. Comparison with cross-validation.

Multi-COVID-Net are twofold: first, the pre-processing step is not required to test the images, since the data augmentation step helps to resize the test images. Second, the pre-trained CNN architecture is optimized using MOGOA on CXR images to avoid overfitting issues to give better performances.

The limitation of the proposed work is as follows:

- Although the proposed Multi objective optimized ensemble deep learning method provides the best performance, but it is computational expensive as it uses two DL networks to improve the efficiency.
- For implementation of multi objective optimization, it is essential that Pareto optimal solutions should be well distributed across all objectives.

5. Conclusions and future work

In this paper, a Multi-COVID-Net model was proposed to diagnose COVID-19 using chest X-ray images. For the diagnosis of COVID-19, two pre-trained CNN architecture such as InceptionV3 and ResNet50 were used. After a comprehensive literature review, it was discussed that the current gap is the lack of parameter tuning of CNN using an optimization algorithm that can substantially improve the accuracy. The MOGOA algorism was then employed for the effective training of these CNN networks as the main contribution. The Multi-COVID-Net model was tested on a wide range of dataset images. The Multi-COVID-Net model performs superior and produced better performance results than other DL networks and meta-heuristic optimization algorithms.

For future work, it is recommended to propose mechanisms for reducing the computational cost of the proposed method, especially in the archiving mechanism. Also, using proper solution selection technique from the archive to improve the distribution of obtained Pareto optimal solutions across all objectives.

Code availability

The developed code for this study will be provided based on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L.J. Carter, L.V. Garner, J.W. Smoot, Y. Li, Q. Zhou, C.J. Saveson, et al., Assay techniques and test development for COVID-19 diagnosis, 2020.
[2] M.Z. Alom, M.M. Rahman, M.S. Nasrin, T.M. Taha, V.K. Asari, Covid_Mtnet: COVID-19 detection with multi-task deep learning approaches, 2020, ArXiv Prepr arXiv:200403747.
[3] M. Alazab, A. Awajan, A. Mesleh, A. Abraham, V. Jatana, S. Alhyari, COVID-19 prediction and detection using deep learning, Int. J. Comput. Inf. Syst. Ind. Manag. Appl. 12 (2020) 168–181.
[4] P.K. Sethy, S.K. Behera, Detection of coronavirus disease (covid-19) based on deep features, 2020, Preprints 2020030300 (2020).
[5] S. Hassantabar, M. Ahmadi, A. Sharifi, Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches, Chaos Solitons Fractals 140 (2020) 110170.
[6] J.E. Iglesias, M.R. Sabuncu, Multi-atlas segmentation of biomedical images: a survey, Med. Image Anal. 24 (2015) 205–219.
[7] T. Hinz, N. Navarro-Guerrero, S. Magg, S. Wermter, Speeding up the hyperparameter optimization of deep convolutional neural networks, Int. J. Comput. Intell. Appl. 17 (2018) 1850008.
[8] J. Zhang, Y. Xie, Y. Li, C. Shen, Y. Xia, Covid-19 screening on chest x-ray images using deep learning based anomaly detection, 2020, ArXiv Prepr arXiv:200312338.
[9] L.O. Hall, R. Paul, D.B. Goldgof, G.M. Goldgof, Finding covid-19 from chest x-rays using deep learning on a small dataset, 2020, ArXiv Prepr arXiv:200402060.
[10] L. Wang, A. Wong, COVID-net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images, 2020, ArXiv Prepr arXiv:200309871.
[11] A. Altan, S. Karasu, Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique, Chaos Solitons Fractals 140 (2020) 110071.
[12] N.N. Das, N. Kumar, M. Kaur, V. Kumar, D. Singh, Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays, IRBM (2020).
[13] A. Imran, I. Posokhova, H.N. Qureshi, U. Masood, S. Riaz, K. Ali, et al., Ai4covid-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app, 2020, ArXiv Prepr arXiv:200401275.
[14] A.I. Khan, J.L. Shah, M.M. Bhat, Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images, Comput. Methods Programs Biomed. (2020) 105581.
[15] T. Mahmud, M.A. Rahman, S.A. Fattah, CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization, Comput. Biol. Med. 122 (2020) 103869.
[16] S. Minaee, R. Kafieh, M. Sonka, S. Yazdani, G.J. Soufi, Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning, 2020, ArXiv Prepr arXiv:200409363.
[17] M. Nour, Z. Cömert, K. Polat, A novel medical diagnosis model for COVID-19 infection detection based on deep features and Bayesian optimization, Appl. Soft. Comput. (2020) 106580.
[18] T. Ozturk, M. Talo, E.A. Yildirim, U.B. Baloglu, O. Yildirim, U.R. Acharya, Automated detection of COVID-19 cases using deep neural networks with X-ray images, Comput. Biol. Med. (2020) 103792.
[19] H. Panwar, P.K. Gupta, M.K. Siddiqui, R. Morales-Menendez, P. Bhardwaj, V. Singh, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-scan images, Chaos Solitons Fractals (2020) 110190.
[20] S. Vaid, R. Kalantar, M. Bhandari, Deep learning COVID-19 detection bias: accuracy through artificial intelligence, Int. Orthop. (2020) 1.
[21] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, 2020, ArXiv Prepr arXiv:200310849.
[22] E.E.-D. Hemdan, M.A. Shouman, M.E. Karar, Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, 2020, ArXiv Prepr arXiv:200311055.
[23] I.D. Apostolopoulos, T.A. Mpesiana, Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks, Phys. Eng. Sci. Med. (2020) 1.
[24] L.K. Hansen, P. Salamon, Neural network ensembles, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990) 993–1001.
[25] A. Ellegård, Darwin and the General Reader: The Reception of Darwin's Theory of Evolution in the British Periodical Press, University of Chicago Press, 1990, pp. 1859–1872.
[26] J. Kennedy, R. Eberhart, Particle swarm optimization, in: Proc. ICNN'95-International Conf. Neural Networks, Vol. 4, (1995) pp. 1942–1948.
[27] S. Mirjalili, A. Lewis, The whale optimization algorithm, Adv. Eng. Softw. 95 (2016) 51–67.
[28] T. Murata, H. Ishibuchi, MOGA: multi-objective genetic algorithms, IEEE Int. Conf. Evol. Comput. 1 (1995) 289–294.
[29] S.Z. Mirjalili, S. Mirjalili, S. Saremi, H. Faris, I. Aljarah, Grasshopper optimization algorithm for multi-objective optimization problems, Appl. Intell. 48 (2018) 805–820.
[30] Chest Imaging, This is a thread of COVID-19 CXR (all SARS-CoV-2 pcr+) from my hospital (Spain), version 1, 2020, Retrieved 06 May 2020 from https://twitter.com/ChestImaging/status/1243928581983670272.
[31] SIRM COVID-19 Database, (2020, MAy) Italian society of medical and interventional radiology COVID-19 dataset, version 1, 2020, Retrieved 05 May 2020, from https://www.sirm.org/ category/senza-categoria/covid-19/.
[32] J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, Covid-19 image data collection: Prospective predictions are the future, 2020, ArXiv Prepr arXiv:200611988.
[33] L. Wang, A. Wong, Z.Q. Lin, J. Lee, P. McInnis, A. Chung, et al., FIgure 1 COVID-19 Chest X-ray Dataset Initiative n.d..
[34] W. Kong, P.P. Agarwal, Chest imaging appearance of COVID-19 infection, Radiol. Cardiothorac. Imaging 2 (2020) e200028.

[35] P. Mooney, Chest x-ray images (pneumonia), 2018, [Online], https://Www.KaggleCom/Paultimothymooney/Chest-Xray-Pneumonia, Tanggal Akses.
[36] A.P. Dhawan, Medical Image Analysis. 31, John Wiley & Sons;, 2011.
[37] Y. Ben-Haim, E. Tom-Tov, A streaming parallel decision tree algorithm, J. Mach. Learn. Res. (2010) 11.
[38] G.E. Batista, M.C. Monard, et al., A study of K-nearest neighbour as an imputation method, His 87 (2002) 48.
[39] M.E. Mavroforakis, S. Theodoridis, A geometric approach to support vector machine (SVM) classification, IEEE Trans. Neural Netw. 17 (2006) 671–682.
[40] I. Rish, et al., An empirical study of the naive Bayes classifier, in: IJCAI 2001 Work. Empir. methods Artif. Intell. 3, 2001, pp. 41–46.
[41] M. Pal, Random forest classifier for remote sensing classification, Int. J. Remote Sens. 26 (2005) 217–222.
[42] R. Murugan, P. Roy, U. Singh, An abnormality detection of retinal fundus images by deep convolutional neural networks, Multimed Tools Appl. (2020) 1–19.
[43] L. Jiang, Z. Ge, Z. Song, Semi-supervised fault classification based on dynamic sparse stacked auto-encoders model, Chemom. Intell. Lab. Syst. 168 (2017) 72–83.
[44] T. Goel, R. Murugan, S. Mirjalili, D.K. Chakrabartty, Optconet: an optimized convolutional neural network for an automatic diagnosis of COVID-19, Appl. Intell. (2020) 1–16.
[45] X. Xia, C. Xu, B. Nan, Inception-v3 for flower classification, in: 2017 2nd Int. Conf. Image, Vis. Comput. 2017, pp. 783–787.
[46] Holland J.H., Genetic algorithms, Sci. Am. 267 (1992) 66–73.
[47] R. Hooke, T.A. Jeeves, Direct SearchSolution of numerical and statistical problems, J. ACM 8 (1961) 212–229.
[48] H. Ishibuchi, N. Tsukamoto, Y. Nojima, Evolutionary many-objective optimization: A short review, in: 2008 IEEE Congr. Evol. Comput. (IEEE World Congr. Comput. Intell. 2008, pp. 2419–2426.