TRACK 2: MEDICAL APPLICATIONS OF MULTIMEDIA

A COVID-19 X-ray image classification model based on an enhanced convolutional neural network and hill climbing algorithms

Ashwini Kumar Pradhan[1] · Debahuti Mishra[1] · Kaberi Das[1] ·
Mohammad S. Obaidat[2,3,4] · Manoj Kumar[5] (iD)

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract

The classification of medical images is significant among researchers and physicians for the early identification and clinical treatment of many disorders. Though, traditional classifiers require more time and effort for feature extraction and reduction from images. To overcome this problem, there is a need for a new deep learning method known as Convolution Neural Network (CNN), which shows the high performance and self-learning capabilities. In this paper,to classify whether a chest X-ray (CXR) image shows pneumonia (Normal) or COVID-19 illness, a test-bed analysis has been carried out between pre-trained CNN models like Visual Geometry Group (VGG-16), VGG-19, Inception version 3 (INV3), Caps Net, DenseNet121, Residual Neural Network with 50 deep layers (ResNet50), Mobile-Net and proposed CNN classifier. It has been observed that, in terms of accuracy, the proposed CNN model appears to be potentially superior to others. Additionally, in order to increase the performance of the CNN classifier, a nature-inspired optimization method known as Hill-Climbing Algorithm based CNN (CNN-HCA) model has been proposed to enhance the CNN model's parameters. The proposed CNN-HCA model performance is tested using a simulation study and contrasted to existing hybridized classifiers like as Particle Swarm Optimization (CNN-PSO) and CNN-Jaya. The proposed CNN-HCA model is compared with peer reviewed works in the same domain. The CXR dataset, which is freely available on the Kaggle repository, was used for all experimental validations. In terms of Receiver Operating Characteristic Curve (ROC), Area Under the ROC Curve (AUC), sensitivity, specificity, F-score, and accuracy, the simulation findings show that the CNN-HCA is possibly superior than existing hybrid approaches. Each method employs a k-fold stratified cross-validation strategy to reduce over-fitting.

Manoj Kumar contributed equally to this work.

✉ Manoj Kumar
   wss.manojkumar@gmail.com

Extended author information available on the last page of the article.

1 Introduction

In December 2019, a pandemic situation raised due to coronavirus (COVID-19) started in place Wuhan, China. Till July-2021, it is not understood completely about COVID-19, and more than 152,534,452 confirmed cases found worldwide according to the World Health Organization (WHO) [24]. The virus Severe Acute Respiratory Syndrome Coronavirus-2 causes the pandemic COVID-19 syndrome (SARS-COV-2).

COVID-19, that has now spread over the world, is believed to have started in a seafood market in Wuhan, where live animals such as frogs, bats, snakes, snails, birds, pigs, and dogs are frequently wholesaled. After examining the COVID-19 genetic structure, it was determined that the sick person may have visited a seafood market in Wuhan City or touched infected birds or animals. In contrast, later surveys revealed that a few cases of people becoming infected even though they had not visited the city of Wuhan. As a result of this investigation, it was determined that the COVID-19-causing virus was capable of community transmission via a human. This new virus is spread from person to person by intimate contact with an infected person or through exposure to respiratory droplets emitted by an infected person during respiratory secretions. This type of infected person's respiratory droplets can enter a healthy human's respiratory system through the nasopharynx during breathing. The virus can survive for at least three hours in the air and up to 72 hours on glass and stainless-steel objects.

The disease Pneumonia is also caused by the illness of the respiratory system. There are lots of moderate symptoms like fever, hard breathing, coughing, chest pain which are common to both pneumonia and COVID-19. To stop the spread of this epidemic, early diagnosis of COVID 19 is a key factor for the researchers who are working in this domain. COVID-19 patients are diagnosed via reverse transcriptase-polymerase chain reaction (RT-PCR) testing through nasopharyngeal or oropharyngeal swabs but it takes more time, expensive, and also tedious.

It's a challenge for the researcher to create an Artificial Intelligence (AI) or Deep Learning (DL) model to assist physicians, so that they will provide better treatment. In a recent paper [3] have used some pre-trained networks such as MobileNet V2, Visual Geometry Group (VGG-19) along with transfer learning for COVID-19 detection on two data-sets. According to the simulation results, the average performance of the VGG-16 model is superior to the MobileNet V2 model for two-class classification. In paper [32], the authors have proposed that CXR photographs be used to diagnose COVID-19 patients instead of computed tomography (CT) scans and magnetic resonance imaging (MRI) [31]. In the last 10 years, DL has obtained a special position in the area of AI for the application of chest X-ray image classification to achieve accurate results [15, 42].

The rest of the paper is organized as follows; Section 2 detailed explanation on CXR image classification in literature review. Section 3, describes the existing methodology associated with the CXR image classification. Concise presentation of the proposed hybrid method is explained in Section 4. Section 5 provides the simulation-based experimental results, discussion, and different analyses. Finally, the conclusion and future scope is discussed in Section 6.

2 Literature survey

Recently, CNNs have obtained popularity in image-based implementation which requires image pre-processing before deploying them to the network [4]. As compared to traditional

Machine Learning (ML) classifiers, CNN produced superior accuracy for CXR images. The CNNs consist of both feature extraction and classification phases in one system. There are several DL pre-trained models with distinct architecture that were trained for image classification and recognition which are publicly available to all the researchers [18, 33, 34, 45]. The different state-of-the-art pre-trained models of DL are VGG-16, VGG-19 [33], ResNet50 (Residual Neural Network with 50 deep layer) [18], INV3 (Inception) [34], MobileNet and DenceNet121 [45].

Researchers are attempting to develop low-cost and quick Computer-Aided Diagnosis (CAD) procedures for CXR [21, 25, 46], CT [22, 35], and so on. But the WHO has motivated individuals to do chest imaging for those having mild symptoms of COVID-19. The crucial obstacle in COVID-19 detection in CXR-based technique is a large number of training parameters needed in existing DL-based models which produces more computation load and over-fitting issues due to limited availability of COVID-19 CXR images.

In paper [43], the authors have applied Support Vector Machine (SVM), VGG-16, and InceptionV3 models in the pneumonia dataset which is divided into three categories: normal, bacterial, and viral pneumonia for classification of algorithms, and reported 96.6% as the best classification accuracy. The other studies like, [19, 20, 29, 40], the author have developed a CNN model with convolution layers, dense blocks, and flatten layers for classification of pneumonia data using the sigmoid function and discovered a success rate of 93.73% in pneumonia using X-ray images. Adam optimizer has been used for optimizing CNN hyper-parameters on ImageNet [16] where epochs vary from 10 to 100 and batch size varies from 8 to 128. In a recent communication, [3] suggested some state-of-the-art CNN networks [8, 9, 11] utilizing the transfer learning techniques to identify COVID-19 and obtained an accuracy of 96.78% in MobileNet architecture which was superior to others. A DarkCOVIDNet model was proposed in [28] by using binary classification (COVID-19 positive, COVID-19 negative) and trinary classification (COVID-19 positive, COVID-19 negative and Pneumonia) and from the simulation study, they found that higher accuracy has been achieved by two classes dataset over three classes.

In another article, CoroNet is a model proposed by the authors [36] for the classification of four classes (COVID-19 vs. bacterial pneumonia vs. viral pneumonia vs. normal) and three classes (COVID vs. pneumonia vs. normal) and observed accuracy of 89.6% and 94% respectively. Despite the fact that the CNN model produces good results for COVID-19 identification and classification, still a lot of hyper-parameters need to be optimized for better improvement [30]. There are several hyper-parameters are associated with the CNN network such as; kernel dimension, network depth, pooling size, stride size, etc. which have a substantial impact on the different performance statistics. Therefore, optimization techniques are needed for finding near-optimal values of those hyper-parameters. Some recent publications have also attempted to use optimization approaches for hyper-parameter selection. To further increase the CNN model's accuracy, [39] have proposed a hybrid model called deep BayesSqueezeNet which uses Bayesian optimization in CNN to tune the hyper-parameter and obtained an accuracy of 98.3% in three classes dataset (Normal, COVID-19, and Pneumonia). The author proposed an optimized CNN model [17] called Optconet using 2700 images and obtained a precision score of 92.8%. In another study [2] have been proposed a Mobilenet CNN model using extricated features. The research article [27], suggested using three predefined CNN models such as; IN-V3, IN-ResNet V2, and ResNet-50 for the classification of the CXR image dataset as COVID-19 or Normal. Minaee et al. [26] employed a transfer learning technique in a recent work in four CNN models

namely ResNet18, ResNet50, SqueezeNet, and DenseNet121 for the binary classification (COVID-19 positive and NON-COVID-19 negative).

So far, there is no work which optimizes every hyper-parameter in designing a CNN architecture. In 2012 [10], have created a CNN model with five convolution layers(CL) and three fully-connected layers(FCL) to train the ImageNet dataset and achieved good results. Several researchers have altered the architecture of the AlexNet, resulting in enhanced accuracy. The author [13] has improved the accuracy by reducing the filter dimension from 11x11 to 7x7 and the stride from 4 to 2 in the AlexNet model's first layer. In another study, [7] have used 19 layers with (3×3) kernel dimension in every convolutional layer to further enhance the accuracy and it was discovered that the depth of the CNN model plays a significant impact in getting higher performance. But the some of the authors have also proved that adding too many layers leads to an increase in training and testing errors [12, 23]. Till now, there is no direct formula for selecting proper hyper-parameters. So, repeating the trial-and-error method is ineffective and time-consuming for bigger data sets [44].

Hence, the optimization algorithm plays a vital role in choosing the proper value of the hyper-parameter of the CNN model. Grid search is a common method for hyper-parameter optimization of CNN but it is comprehensive because it attempts to combine all possible hyper-parameters. Recently, random search [5] produced better results than a grid search. In recent work, the authors have used Bayesian optimization (BO) [37] to optimize nine hyper-parameters such as epochs, learning rate, weights of CNN, and parameters of contrast normalization of a CNN model.The authors of [44] have employed genetic algorithms to improve two parameters (filter dimension and feature map of convolution layer) of a CNN network and obtained 75% accuracy on CIFER 10 data-sets.

In this research, an attempt has been made to design an efficient classifier for CXR dataset to automate the detection of COVID-19. To enhance the performance of the CNN architecture, a CNN-based optimized hybrid model with an appropriate layer is proposed for fine-tuning the model's parameters using bio-inspired optimization techniques.

3 Materials and methods

The methodology comprises three sub-stages. (a) Controlling over-fitting with stratified k-fold cross-validation; (b) Design of CNN classifier and; (c) Employ PSO, Jaya, and HCA to adjust the parameters of CNN.

3.1 k-fold cross-validation

The k-fold cross-validation method is used to prevent over-fitting. The complete dataset is randomly divided into k-fold partitions , one of which is necessary for testing while the remaining folds are used for training. Finally, the average error produced among all sets is computed. This procedure is repeated unless all data folds have been tested.

3.2 CNN methodology

CNN model was developed by a French computer scientist Yann Andre Lecun and his team in 1998 using the field of computer vision and machine learning. The CNN-based model is extensively utilized in the medical image classification job. Hugh number of public image data-sets such as ImageNet [14], Places [6] are required for the training of non-pre-trained DL models. CNN consists of two parts: (a) feature extraction (b) classification.

Feature extraction comprises CL and pooling layer (PL). The weights and biases of CL and FCL have been tuned by stochastic gradient descent (SGD) algorithm which minimizes the cost function through the backward direction. Through the ReLU activation function, the extracted features are mapped into a new feature space. Between CL and ReLU, the batch normalisation layer (BNL) is used to normalize the gradients and activation function.However max pooling is superior to others due to its faster convergence rate and well generalization [38]. FCL is the final layer of the CNN model, and it classifies the extracted visual feature into a specific class. Figure 1 exhibits a general architecture of a CNN with CL, PL, FCL, and output layer.

If the input size is $(a \times a)$, kernel dimension $(b \times b)$, and stride (c), then the size of the feature matrix (FM)(f) is computed as per (1).

$$f = \lfloor \frac{a - b}{c} \rfloor + 1 \tag{1}$$

If the PL splits the input to small blocks of pixels of size $(d \times d)$, then the pooling layer's output size (p) is calculated using (2).

$$p = \lfloor \frac{f}{d} \rfloor + 1 \tag{2}$$

In this work, two parameters of the CNN model have been optimized such as; filter dimension and the total number of neurons in the first dense layer by HCA optimization techniques while other parameters have been fixed.

3.3 PSO, Jaya and hill climbing algorithm

PSO (Particle Swarm Optimization) is developed by Eberhart and Kennedy which uses the social behavior of bird's flock. Lots of particle makes a swarm and each particle moves in the solution space. Each particle represents a feasible solution in solution space. Due to its faster convergence, it is considered a promising optimization technique.

JAYA algorithm is a recent population-based algorithm which is proposed by Rao in 2016. Due to its impressive characteristics and simplicity, it is easy to use. In its initial search, there is no derivative information. It has only two parameters such as; population size and iteration numbers.

In this work, a new optimization algorithm has been used known as hill climbing which provides optimized results. It is a local search optimization technique that explores the superior solution among neighborhoods after evaluating the present state. It updates the current state until reaches the goal state.
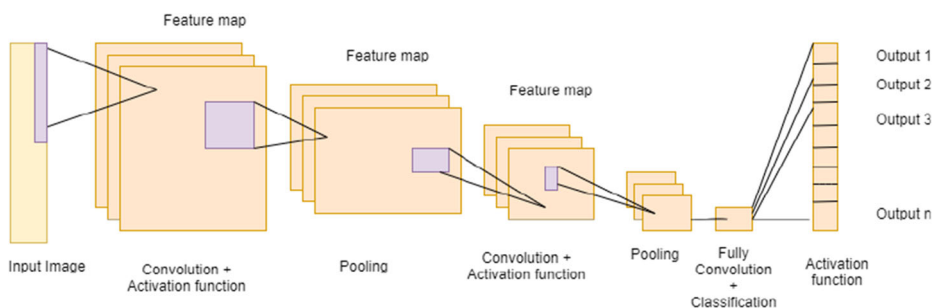
Fig. 1 Block diagram of CNN

4 Proposed CNN-HCA methodology

This section describes the design of the proposed hybrid classifier. Figure 2 represents the overall structure of our proposed hybrid model "CNN-HCA" which classifies a CXR image into either "COVID-19", or "pneumonia".

4.1 Phase 1 (data preparation)

To balance the dataset, 50% of images have been taken for both positive category (110 images) and negative category (110 images) respectively. The balanced dataset has been divided into two phases such as; Training (80%) and Testing (20%). Prior to applying the training and testing phases, several data augmentation techniques such as horizontal and vertical flip, zooming, width shift range, rotation and width height were already applied in the dataset. Due to variation in dimension of CXR images, all were resizing to 224 × 224.

4.2 Phase 2 (selection of hyper-parameters)

To create a new CNN classifier, minor changes of some hyper-parameters is required. For optimization, two hyper-parameters have considered in this work: filter size and the total number of neurons in the first dense layer.

4.3 Phase 3 (learning phase)

For feature extraction, the original existing CNN model is replaced with a new one which consists of 9 layers such as; three CL, three max-pooling layers, one flatten layer, and two
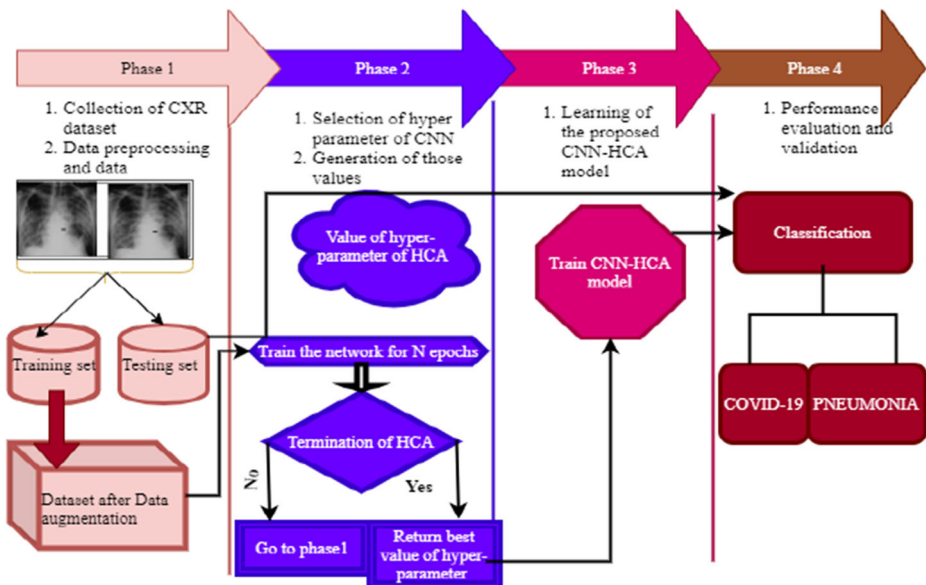

Fig. 2 COVID-19 CXR image classification block diagram of the proposed hybrid model "CNN-HCA"

dense layers. HCA determines the total number of neurons in the first dense layer and appropriate filter size which uses the RELU activation function. The new classifier learns for 20 epochs before beginning the fine-tuning procedure.

4.4 Design of CNN model

Convolution, relu, dropout, pooling, and dense layers are one of the layers in the CNN model. Each layer is significant in its own right. Despite the fact that CNNs have shown promise in a variety of classification issues, determining the appropriate CNN configuration for a certain application is far from basic. The presented configurations of CNN in the literature are largely based on trial and error or are influenced by previous relevant studies. The initial CNN structure is developed in our situation by referring to the published articles.

The proposed CNN model's architecture in this study consists of three CL with 16, 32, and 64 filters, as well as two dense layers with 512 and 1 neuron. A (3×3) filter dimension has been considered for the architecture as well as a dropout rate of 0.5. Excluding the last CL, a 2×2 maximum pooling was considered for each layer. A (1×1) max-pooling was used in the last convolutional layer. The architectural property of the proposed CNN is represented in Table 1. The proposed network provides approximately 57,377 parameters as compared to other models. A 5-fold cross-validation was used in the experiment. In this research, (240×240) CXR images were resized to (15×15).

To convert the features obtained from the convolution block into one-dimensional features, we employ FCL, which has three layers: flatten, dropout, and dense. In our proposed method, we fix dropout to 0.5 and set the number of the last dense layer neuron to 1. The sigmoid layer is used to classify the features taken from the FCL. The number of units for the SoftMax layer is determined by the number of categories.

The pseudo-code of the "CNN-HCA" is represented in Algorithm 1 which improves the quality of the solution. At each iteration of optimization methods, it updates the dimension of filter size along with the number of neurons in the first dense layer in the model structure. The proposed algorithm is repeated until the final condition is not satisfied.

Note: [Condition for termination of HCA]: Either the iteration number reaches 50 or the fitness of particle doesn't change for 5 generations continuously.

4.5 Phase 4 (performance evaluation phase)

To analyse the suggested model, five metrics were used: classification accuracy, sensitivity, specificity, ROC, and AUC. In this study, HCA is used to reduce the misclassification rate, which is specified as a cost function in (3).

$$f = min(1 - Accuracy) \tag{3}$$

Here, Accuracy is calculated using (4)

$$Accuracy = \frac{\sum_{k=1}^{n} \sum_{l=1}^{c} A(k,l).B(k,l)}{n} \tag{4}$$

where n and c represents the total number of instances and classes respectively.

- $A(k,l) = 1$, If an instance k is of class l else 0.
- $B(k,l) = 1$, , if the predicted class of instance k is l else 0.

The following are the study's primary contributions:

1: **INPUT:** Set of the pre-defined building block of CNN, Population size (N), Maximum iteration, Image dataset for classification, Number of epochs for both particle evaluation, and best CNN architecture

2: **OUTPUT:** The best architecture of CNN to classify the COVID-19 CXR image dataset.

3: Initialization of particles

4: Compute fitness evaluation f of all particles using a cross-entropy loss function

5: Initialize bounds of hyper-parameter of CNN i.e filter size and number of neurons

6: Assume s = starting solution

7: **for** iteration = 1 to Maximum iteration **do**

8:      **for** i = 1 to N **do**

9:          **while** $f(k) \Leftarrow f(s)$ **do**

10:             **for** all k ε - neighbours(s) **do**

11:                 Generates an $k \Leftarrow \varepsilon$ - neighbour(s)

12:                 **if** fitness(k) > fitness(s) **then**

13:                     Replace k with s;

14:                 **end if**

15:             **end for**

16:         **end while**

17:     **end for**

18: **end for**

Algorithm 1 Pseudo-code for CNN-HCA algorithm.

- The proposed method has converged at the 42nd iteration with a very less error rate of 0.11 as compared to other hybridized models to generalize the classification of CXR images and show the best fit for the dataset.
- The suggested hybrid classifier "CNN-HCA" is intended to divide CXR images into two categories.
- Selection of bio-inspired technique which is being able to optimize the CNN parameters such as; Filter dimension and the total number of neurons in the first dense layer concerning improvement in classification performance.
- A comparative study is carried out between the proposed CNN model and pre-trained CNN models.

Table 1 Architectural properties of the proposed CNN

Layer	Layer Type	Input	Filters	Filter/ Kernel Dimension	Output	Parameters	Activation
Layer 1	CL 1	$15x15x3$	16	$3x3$	$15x15x16$	448	ReLU
Layer 2	MPL1	$15x15x16$		$2x2$	$15x15x16$	0	
Layer 3	CL 2	$7x7x16$	32	$3x3$	$7x7x32$	4640	ReLU
Layer 4	MPL2	$7x7x32$		$2x2$	$3x3x32$	0	
Layer 5	CL 3	$3x3x32$	64	$3x3$	$3x3x64$	48496	ReLU
Layer 6	MPL 3	$3x3x64$		$2x2$	$1x1x64$	0	
Layer 7	Flatten	$1x1x64$			64		
Layer 8	Dense 1			$1x1$	512	33280	ReLU
Layer 9	Dense 2			$1x1$	1	513	Sigmoid

- For adjusting the hyper-parameters of CNN, three optimization algorithms were used: HCA, Jaya, and PSO, and the suggested technique discovered the best optimum values of the hyper-parameters.

5 Results and discussion

5.1 Description of the data set

The data can be found on the Kaggle website which holds two categories of CXR image folders as; (a) COVID-19 (b) Pneumonia. Before applying to the training and testing phase, several data augmentation techniques such as horizontal and vertical flip, zooming, width shift range, rotation and width height were already applied in the dataset. Not only does it improve CNN performance, but also prevents the over-fitting issue. Table 2 explains the different parameters and the values used in the dataset.

The composition of the dataset is narrated in Table 3. There are 220 numbers of CXR images were selected for this work and divided into 80:20 ratios, where 80%, 20% of data were used for both training testing respectively. The dataset is balanced and consists of 50% of images positive category (110 images) and 50% of images in are negative category (110 images). There are 120 male patients and 65 female patients. The average age of the COVID-19 cluster is 55.5 years. The training and testing dataset comprise 176 and 44 number of CXR images respectively. Due to variation in dimension in images, all were resizing to 224×224. Different types of CXR images from the dataset have been represented in Fig. 3. The COVID-19 CXR image (Fig. 3(a)) manifests with diffuses of the lung due to whiteness while the Pneumonia CXR image (Fig. 3(b)) depicts an area of lung inflammation and focal lobar consolidation.

5.2 Hardware and software setup

For CXR image classification, an ordinary high-performance computer like 8GB memory, Intel i7 (2.3 GHz) Processor, 256GB solid-state drive disk, NVIDIA Tesla K80 was used. In this paper, a Google Cloud GPU was used to accelerate the process also. A python program was used to test the classification result while needs a virtual machine in the GPU cloud since CNN techniques require high computing.

5.3 Model evaluation criteria

In this work $Accuracy$, $sensitivity$, $specificity$, $f - score$, ROC, and AUC have been taken into consideration for the performance of the model for COVID-19/Pneumonia experiments. Model having higher ROC and AUC value is more proficient and effective in differentiating both the patients. Equations (5), (6), (7), (8), and (9) represent the $Accuracy$, $sensitivity$, $specificity$, $precision$, and $f - score$ respectively.

$$Accuracy = (tp + tn) / (tp + tn + fp + fn) \tag{5}$$

$$sensitivity = \frac{tp}{tp + fn} \tag{6}$$

$$specificity = \frac{tn}{tn + fp} \tag{7}$$

Table 2 Data augmentation parameters utilised in the dataset before the training and testing phases

Parameters	Values
Horizontal flip	TRUE
Vertical flip	TRUE
Rotation	3 degrees
Zoom in	0.06
Width shift range	0.07
Width height range	0.06

$$precision = \frac{tp}{tp + fp} \tag{8}$$

$$f - score = \frac{2 \times t\,p}{2 \times t\,p + fp + fn} \tag{9}$$

Where tp, tn, fp, fn represent true positive, true negative, false-positive, and false-negative respectively. f-score comprises the harmonic mean of recall and precision values of both classes.

5.4 Experimental results

Table 4 represents a summary of the comparison of performance in terms of accuracy between pre-trained CNN models and the proposed CNN model using the same CXR COVID dataset. The parameters of different pre-trained deep learning models are explained below.

In VGG-16 network, 16 means it contains 16 layers with different weights. This network is rather huge, with over 138 million (approximately) parameters. The most distinctive feature of VGG16 is that, rather than having a huge number of hyper-parameters, they worked on having 3x3 filter convolution layers with a stride 1 and always utilised the same padding and max-pool layer of 2x2 filter stride 2. Throughout the design, the convolution and max pool layers are arranged in the same way. Finally, there are two FC layers and a soft-max for output.

VGG19 is a VGG variation with 19 layers (16 convolution layers, 3 FC layer) whereas 5 Max-Pool layers and 1 SoftMax layer not taken into account. Capsule Networks (CapsNet) are networks that can retrieve geographical information in order to avoid the data loss that occurs during pooling operations. ResNet50 is a ResNet variation of 48 Convolution layers, 1 Max-Pool layer, and 1 Average Pool layer. There are 3.8 x 109 floating point operations in it. DenseNet-121 has One 7x7 Convolution, 58 3x3 Convolution, 61 1x1 Convolution, 4 Average-pool, and 1 FC layer. MobileNet employs depth-wise separable convolutions.

It's observed from Table 4, the proposed 9-layer CNN model produces 86% accuracy which is superior to all the pre-trained architectures.

The effect of several hyper-parameters such as the loss function, the number of epochs, and the batch size on the accuracy of many architectures of DL on same CXR dataset by the

Table 3 Data-set description

Sample Type	Training Data-set		Testing Data-set		Sample Source
	#Samples Positive	#Samples Negative	#Samples Positive	#Samples Negative	
COVID-19	56	56	14	14	Kaggle CXR dataset
Pneumonia	32	32	8	8	
Total	176 (100%)		44 (100%)		

(a) COVID-19   (b) Pneumonia
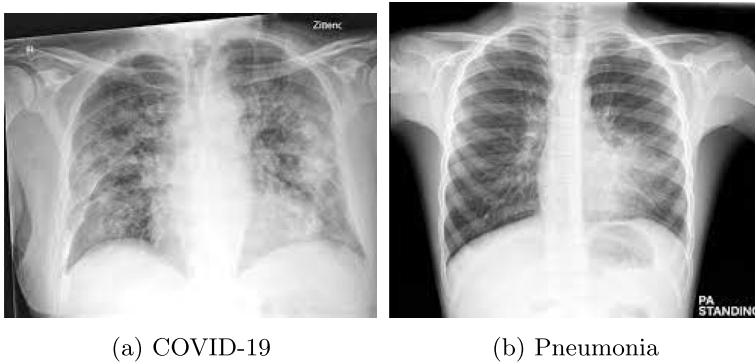
Fig. 3 Different samples of CXR image

trial-and-error approach has also been observed from Table 4. The CNN model provides the best results as compared to other pre-trained CNN models using cross-entropy loss function, 20 epochs, and a batch size of 10 in our proposed CNN model accomplishes the best results as compared with other models.

To prevent the over fitting issue, different explorations results of the proposed CNN model with the last unfrozen ConvLayer have been shown in Table 5. As the final accuracy of the CNN model does not depend on one factor, so a fair comparison of different parameters is needed. From the analysis, it can be found that the highest accuracy of the proposed CNN model can be obtained with the learning rate of 0.001, a dropout rate of 0.5, with batch BNL.

5.5 Setting up the HCA for the hyper-parameters selection

The searching space for filter size and the total number of neurons in the first FCL are [3,10] and [50,600] respectively where the maximum number of iteration and population are chosen to 50 and 30, respectively, represented in Table 6. The results of the experiment showed that when the number of epochs is greater than twenty, the training process for each of the HCA requires exponential time, and when the number of epochs is less than twenty, the proposed model's results are not sufficiently accurate. As a result, the number of epochs was limited to twenty. The purpose of using the HCA is to lower the testing set's loss rate as much as feasible.

The best settings for kernel size and the number of neurons in the first dense layer were identified after approximately 7 hours of training. The optimal values for the hyperparameters chosen by HCA are shown in Table 7 where filter size and the number of neurons are 7, 512 respectively.

5.6 Convergence analysis

As shown by Fig. 4 and Table 8, it has been observed that the proposed CNN-HCA hybrid model has achieved the highest accuracy of 90% (90% of the test set correctly), sensitivity of 92%, specificity of 89.10%, F-score of 90.25% and AUC value of 97.94% on two-class

Table 4 Comparison of accuracy between CNN model and pre-trained deep learning models using the different combination of the epoch, batch size, loss function in the same CXR COVID dataset by manual search to quantify hyper-parameter values

Pre-trained model	Epoch	Batch Size	Loss function	Accuracy
VGG-16	10	5	Cross Entropy	0.78
	20	10		0.75
	30	20		0.79
VGG-19	10	5	Label Smoothing	0.8
	20	10		0.78
	30	20		0.76
INV3	10	5	Label Smoothing	0.84
	20	10		0.79
	30	20		0.65
Caps Net	10	5	Cross Entropy	0.8
	20	10		0.78
	30	20		0.71
DenseNet121	10	5	Label Smoothing	0.69
	20	10		0.79
	30	20		0.81
ResNet50	10	5	Cross Entropy	0.81
	20	10		0.76
	30	20		0.82
MobileNet	10	5	Label Smoothing	0.65
	20	10		0.74
	30	20		0.78
Proposed CNN model	10	5	Cross Entropy	0.82
	20	10		**0.86**
	30	20		0.81

The bold value indicates the best results

Table 5 Estimation of dropout, batch normalization layer, and learning rate for proposed CNN model in the final Convolution layer

Learning rate	Dropout	BNL	Accuracy
0.001	0.7	Yes	0.77
0.001	0.9	No	0.74
0.0009	0.3	No	0.82
0.0005	0.4	No	0.81
0.001	0.5	Yes	0.82
0.0001	0.5	Yes	0.86

Table 6 Necessary parameter values for HCA in collaboration with the suggested CNN

Parameters	Values
Highest iterations	50
Population	30
Filter size bounds	[3,10]
Neurons in first FCL bounds	[50,600]

Table 7 HCA calculated the optimal settings for the hyper-parameters

| Hyper-parameters	Best feasible value
Filter dimension	7
Total number of neurons of the first FCL	512

classification over the remainder of the models including CNN-Jaya (87.95%), and CNN-PSO (88.94%) in terms of accuracy. Further, the F- score for the CNN-HCA model (90.25%) is higher than CNN-Jaya (87.21%) and CNN-PSO (88.42%). To make a fair comparison, we fixed the number of iterations to 50 in all the hybrid models.

From Fig. 5b, it is also found that our proposed method has converged at the 42nd iteration with a very less error rate as compared to other hybridized models to generalize the classification of CXR images and show best-fit for the dataset. In comparison to other hybrid models, the results in Table 9 revealed that the proposed technique is more dependable, valid, and superior.

It has observed from Table 8 that the suggested CNN-HCA model outperformed the previous hybrid architecture models in terms of accuracy, sensitivity, specificity, F-score, and AUC, with 90%, 92%, 89.1%, 90.25%, and 97.94% respectively. The other approaches [1, 6, 41] were selected for comparison with the proposed approach for classification of COVID-19 using the same CXR images and obtained a superior result for the proposed model as shown in Table 9.

The dataset used in our proposed model is balanced, whereas the dataset in earlier work is imbalanced. Joy Iong-Zong Chen [6] has made the Cohen's dataset, which comprises of 60k around pictures with 400 positive COVID-19 X-ray. In classifying COVID-19, the suggested CNN model with 5 convolutional layers obtained an accuracy of 85%, 69.5% precision, and 80.5% recall. To classify COVID-19, Alkas et al. [1] developed an LSTM network in another work. The experimental findings demonstrate that the proposed model has an accuracy of 86.66 %, F1-score of 91.89%, a precision of 86.75%, a recall of 99.42%, and an AUC of 62.50% in identifying patients with COVID-19 illness.

Xiangjun et al. [41] acquired chest CT pictures from 495 patients at three Chinese hospitals. They used a deep network (ResNet-50) to train a multi-view hybrid approach to identify patients with COVID-19 and obtained AUC, accuracy, sensitivity, and specificity of

(a) sensitivity, specificity, F-score　　　　(b) Accuracy
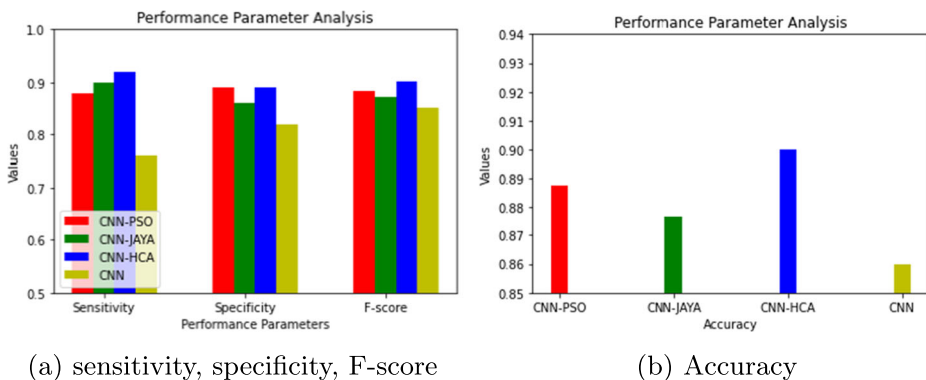
Fig. 4 Comparison of CXR image dataset with other hybridized models

Table 8 Comparison of the suggested hybrid model's sensitivity, specificity, F-score, AUC, and accuracy to others

Models	Sensitivity	Specificity	F-Score	AUC	Accuracy
CNN	0.76	0.82	0.8511	-	0.86
CNN-PSO	0.88	0.89	0.8842	0.9204	0.8894
CNN-Jaya	0.9	0.86	0.8721	0.9028	0.8795
CNN-HCA	0.92	0.891	0.9025	0.9794	0.9

0.819, 0.760, 0.811, and 0.615 in the test dataset respectively. However, on two-class classification of the same CXR dataset with different samples, the proposed CNN-HCA hybrid model achieved the highest accuracy of 90% (on test set), sensitivity of 92%, specificity of 89.10%, F-score of 90.25%, and AUC value of 97.94%.

6 Conclusion

Using the CXR dataset, this study proposed an effective hybrid classification approach. The suggested CNN-HCA model beats existing competitive hybrid ML models, according to a comprehensive simulation analysis. The filter size and the total number of neurons in first FCL were optimised using a specific meta-heuristic optimization algorithm called HCA. Due to the facility of automatic feature extraction, CNN-based optimization techniques are better than conventional methods. People inflamed with COVID-19 are probably to go through everlasting harm within the lungs, which may later initiate death. This takes a look and is aimed to differentiate humans with broken lungs as a result of COVID-19 from ordinary people or pneumonia. The detection of COVID-19 changed into the complete use of deep learning models. Since it's vital to discover COVID-19 that unfold hastily and globally, AI strategies are used to carry out this correctly and quickly. The proposed model aims to provide quick and correct results.

The proposed approach is only suitable for two hyper parameter such as; kernel size and number of neurons in first Dense layer which is the limitation of work. If we will add more hyper-parameters such as training size, training rate, and activation functions in the process

(a) RoC　　　　　　　　　　　　(b) Convergence Curve
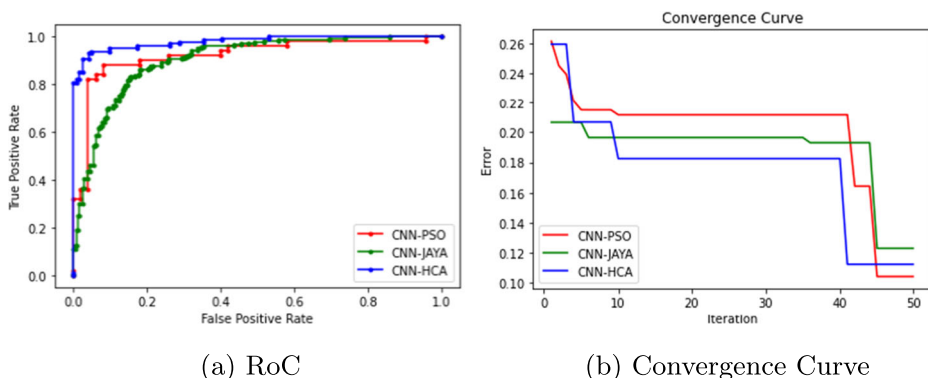
Fig. 5 Comparison of hybridized models based on RoC and Convergence Curve

Table 9 A comparison of the proposed approach's results with the best results obtained by the other offered techniques

Approaches	# samples	# Classes	accuracy	precision	recall	f-Score
CNN [6]	400 COVID-19(+), 60 COVID-19(-)	2 Classes	85%	69.5%	80.5%	-
LSTM [1]	520 COVID-19(-), 80 Healthy(+)	2 Classes	86.66%	86.7%	99.42%	91.8%
ResNet-50 [41]	195 COVID-19(+), 258 COVID-19 (-)	2 Classes	76.00%	61..5%	91%	-
Proposed study	110 COVID-19(+),110 Pneumonia	2 Classes	90%	94%	92%	90.25%

of optimization, the proposed model does not provide well performance in terms of accuracy (Less than 80% approximate). Selection of CNN layers and image enhancement algorithms are the threats to the internal validity of this design. If we will increase or decrease the number of CNN layers, then the performance of experimental result may vary. The effect of image enhancement algorithms will also impact on the performance of CNN models in deep learning and transfer learning.

Its been observed from literature survey, the HCA has been used more for feature extraction problem, weight initialization of CNN, and training of CNN, but the authors have not applied for tuning the hyper-parameter of CNN in in the domain of COVID-19 image classification. So, for the first time it has been applied for tuning the hyper-parameters of CNN in application to COVID-19 classification. PSO and Jaya have adopted hyper-parameters in a narrow sense including only the number of layers without changing the overall architecture but HCA has taken into it in a broader sense by considering the kernel size and number of neurons in first FCL in this work.

The proposed model is more efficient in terms of network size (only 9 number of layers) as compared to any pre-trained model trained for more than one purpose. In case of VGG16 it is trained for 1000 different categories and even if 2 categories, to train the whole model again, because layers can't change the deep knowledge of VGG16.

The proposed model can be utilized in other domains with binary and multiclass datasets. The main advantage of CNN compared to its predecessors is that it automatically detects the important features without any human supervision. Also, another key feature is that deep convolutional networks are flexible and work well on image data. Convolutional layers exploit the fact that an interesting pattern can occur in any region of the image. There are many types of HCA are available such as simple, Steepest-Ascent, Stochastic. But the authors have used simple HCA in this work which required limited computational power as compared to other optimization techniques.

The best accuracy may not be attained if the dimension of the input photos in the dataset varies. So, regardless of the resize parameter, dealing with very low-resolution photographs remains a barrier for the suggested approach. The best accuracy may not be attained if the dimension of the input photos in the dataset varies. So, regardless of the resize parameter, dealing with very low-resolution photographs remains a barrier for the suggested approach to the future. The number of data samples utilized to learn the suggested approach could be increased in the future to improve its overall performance in the analysis of COVID-19. Furthermore, a wide range of illnesses that cause pneumonia can be accelerated, and the proposed method can be utilised to distinguish them from COVID-19.

Declarations

Competing interests No competent interest applicable.

Conflict of Interests The authors have no conflict of interest.

References

1. Alakus TB, Turkoglu I (2020) Comparison of deep learning approaches to predict COVID-19 infection. Chaos Solitons Fract 140:110120
2. Apostolopoulos ID, Aznaouridis SI, Tzani MA (2020) Extracting possibly representative COVID-19 biomarkers from X-ray images with deep learning approach and image data related to pulmonary diseases. Journal of Medical and Biological Engineering 1
3. Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640
4. Ayyar MP, Benois-Pineau J, Zemmari A (2021) A hierarchical classification system for the detection of covid-19 from chest X-Ray images. In: Proceedings of the IEEE/CVF international conference on computer vision, pp 519–528
5. Bergstra J, Bengio Y (2012) Random search for hyper-parameter optimization. J Machine Learn Res 13(2)
6. Chen JIZ (2021) Design of accurate classification of COVID-19 disease in X-ray images using deep learning approach. Journal of ISMAC 3(02):132–148
7. Chithaluru P, Al-Turjman F, Kumar M, Stephan T (2020) I-AREOR: An energy-balanced clustering protocol for implementing green IoT in smart cities. Sustainable Cities Soc 61:102254
8. Chithaluru P, Al-Turjman F, Stephan T, Kumar M, Mostarda L (2021) Energy-efficient blockchain implementation for cognitive wireless communication networks (CWCNs). Energy Rep 7:8277–8286
9. Chithaluru P, Fadi AT, Kumar M, Stephan T (2021) MTCEE-LLN: Multi-layer threshold cluster-based energy efficient low power and lossy networks for industrial internet of things. IEEE Internet of Things Journal
10. Chithaluru PK, Khan MS, Kumar M, Stephan T (2021) ETH-LEACH: An Energy enhanced threshold routing protocol for WSNs. Int J Commun Syst e4881
11. Chithaluru P, Kumar S, Singh A, Benslimane A, Jangir S. K. (2021) An energy-efficient routing scheduling based on fuzzy ranking scheme for internet of things (IoT). IEEE Internet of Things Journal
12. Chithaluru P, Tiwari R, Kumar K (2019) AREOR–Adaptive ranking based energy efficient opportunistic routing scheme in Wireless Sensor Network. Comput Netw 162:106863
13. Chithaluru P, Tiwari R, Kumar K (2021) ARIOR: Adaptive ranking based improved opportunistic routing in wireless sensor networks. Wirel Pers Commun 116(1):153–176
14. Deng J, Dong W, Socher R, Li LJ, Li K, Fei-Fei L (2009) Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. Ieee, pp 248–255
15. Deng X, Zhang Y, Yang S, Tan P, Chang L, Yuan Y, Wang H (2017) Joint hand detection and rotation estimation using CNN. IEEE Trans Image Process 27(4):1888–1900
16. Glorot X, Bengio Y (2010) Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics. JMLR Workshop and Conference Proceedings, pp 249–256
17. Goel T, Murugan R, Mirjalili S, Chakrabartty DK (2021) Optconet: an optimized convolutional neural network for an automatic diagnosis of COVID-19. Appl Intell 51(3):1351–1366
18. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
19. Heidari M, Mirniaharikandehei S, Khuzani AZ, Danala G, Qiu Y, Zheng B (2020) Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. Int J Med Inform 144:104284
20. Hu T, Khishe M, Mohammadi M, Parvizi GR, Karim SHT, Rashid TA (2021) Real–time COVID-19 diagnosis from X-Ray images using deep CNN and extreme learning machines stabilized by chimp optimization algorithm. Biomed Signal Process Control 68:102764
21. Khan AI, Shah JL, Bhat MM (2020) Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. Comput Methods Programs Biomed 196:105581

22. Ko H, Chung H, Kang WS, Kim KW, Shin Y, Kang SJ, Lee JH, Kim YJ, Kim NY, Jung H, Lee J (2020) COVID-19 Pneumonia diagnosis using a simple 2D deep learning framework with a single chest CT image: model development and validation. J Med Int Res 22(6):e19569

23. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. Adv Neural Inform Process Syst 25:1097–1105

24. Kumar D, Malviya R, Sharma PK (2020) Corona virus: a review of COVID-19. EJMO 4(1):8–25

25. Loey M, Smarandache F, M Khalifa NE (2020) Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning. Symmetry 12(4):651

26. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ (2020) Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. Med Image Anal 65:101794

27. Narin A, Kaya C, Pamuk Z (2021) Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. Pattern Anal Applic 1–14

28. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med 121:103792

29. Pathan S, Siddalingaswamy PC, Ali T (2021) Automated Detection of Covid-19 from Chest X-ray scans using an optimized CNN architecture. Appl Soft Comput 104:107238

30. Pradhan AK, Das K, Mishra D, Mishra S (2021) Exploration of hyperparameter in extreme learning machine for brain mri datasets. In: Intelligent and Cloud Computing. Springer, Singapore, pp 449–457

31. Pradhan A, Mishra D, Das K, Panda G, Kumar S, Zymbler M (2021) On the classification of MR images using "ELM-SSA" coated hybrid model. Mathematics 9(17):2095

32. Rubin GD, Ryerson CJ, Haramati LB, Sverzellati N, Kanne JP, Raoof S, Schluger NW, Volpi A, Yim JJ, Martin IB, Anderson DJ (2020) The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the Fleischner Society. Radiology 296(1):172–180

33. Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC (2015) Imagenet large scale visual recognition challenge. Int J Comput Vision 115(3):211–252

34. Sheng T, Feng C, Zhuo S, Zhang X, Shen L, Aleksic M (2018) A quantization-friendly separable convolution for mobilenets. In: 2018 1st workshop on energy efficient machine learning and cognitive computing for embedded applications (EMC2). IEEE, pp 14–18

35. Singh D, Kumar V, Kaur M (2020) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution–based convolutional neural networks. European Journal of Clinical Microbiology & Infectious Diseases 39(7):1379–1389

36. Shukla A, Ramdasani U, Vinzuda G, Obaidat MS, Tanwar S, Kumar N (2021) Bcovx: Blockchain-based covid diagnosis scheme using chest x-ray for isolated location. In: ICC 2021-IEEE international conference on communications. IEEE, pp 1–6

37. Snoek J, Larochelle H, Adams RP (2012) Practical bayesian optimization of machine learning algorithms. Advances in neural information processing systems 25

38. Skalak DB (1994) Prototype and feature selection by sampling and random mutation hill climbing algorithms. In: Machine Learning Proceedings 1994. Morgan Kaufmann, pp 293–301

39. Ucar F, Korkmaz D (2020) COVIDIagnosis-net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. Med Hypotheses 140:109761

40. Varshni D, Thakral K, Agarwal L, Nijhawan R, Mittal A (2019) Pneumonia detection using CNN based feature extraction. In: 2019 IEEE international conference on electrical, computer and communication technologies (ICECCT). IEEE, pp 1–7

41. Wu X, Hui H, Niu M, Li L, Wang L, He B, Yang X, Li L, Li H, Tian J, Zha Y (2020) Deep learning-based multi-view fusion model for screening 2019 novel coronavirus pneumonia: a multicentre study. Europ J Radiol 128:109041

42. Wu X, Jiang G, Wang X, Xie P, Li X (2019) A multi-level-denoising autoencoder approach for wind turbine fault detection. IEEE Access 7:59376–59387

43. Yadav SS, Jadhav SM (2019) Deep convolutional neural network based medical image classification for disease diagnosis. Journal of Big Data 6(1):1–18

44. Young SR, Rose DC, Karnowski TP, Lim SH, Patton RM (2015) Optimizing deep learning hyper-parameters through an evolutionary algorithm. In: Proceedings of the workshop on machine learning in high-performance computing environments, pp 1–5

45. Zhu X, Mao Z, Chen Z, Li Y, Wang Z, Wang B (2021) Object-difference drived graph convolutional networks for visual question answering. Multimed Tools Appl 80(11):16247–16265

46. Zhang HT, Zhang JS, Zhang HH, Nan YD, Zhao Y, Fu EQ, Xie YH, Liu W, Li WP, Zhang HJ, Jiang H (2020) Automated detection and quantification of COVID-19 pneumonia: CT imaging analysis by a deep learning-based software. Eur J Nucl Med Mol Imaging 47(11):2525–2532

Affiliations

Ashwini Kumar Pradhan[1] · Debahuti Mishra[1] · Kaberi Das[1] · Mohammad S. Obaidat[2,3,4] · Manoj Kumar[5] 🆔

Ashwini Kumar Pradhan
ashwini.10.pradhan@gmail.com

Debahuti Mishra
debahutimishra@soa.ac.in

Kaberi Das
kaberidas@soa.ac.in

Mohammad S. Obaidat
msobaidat@gmail.com

[1] Department of Computer Science and Engineering, Siksha O Anusandhan (Deemed to Be University), Khandagiri, Bhubaneswar, 751030, Odisha, India

[2] Distingsuhed Professor, Indian Institute of Technology, Dhanbad, India

[3] KASIT, University of Jordan, Amman, Jordan

[4] University of Science and Technology Beijing, Beijing, China

[5] Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, Dubai Knowledge Park, Dubai, United Arab Emirates