Combining the advantages of AlexNet convolutional deep neural network optimized with anopheles search algorithm based feature extraction and random forest classifier for COVID-19 classification

Sumaiya Begum Akbar¹ | Kalaiselvi Thanupillai² | Suganthi Sundararaj³

¹Department of Electronics and Communication Engineering, R.M.D Engineering College, Chennai, India
²Department of Electronics and Instrumentation Engineering, Easwari Engineering College, Chennai, India
³Department of Computer and communication, Sri Sairam Institute of Technology, Chennai, India

Abstract

In this article, COVID-19 detection and classification framework based on anopheles search optimized AlexNet convolutional deep neural network for random forest classifier is implemented. Here, the COVID-19 dataset is taken from Joseph Paul Cohen database. Then, the input images are preprocessed with the help of fuzzy gray level difference histogram equalization technique (FGLHE) and fuzzy stacking technique for color enhancement and noise elimination in the input images. The FGLHE technique and fuzzy stacking technique are combined together and forms into stacked dataset image. This stacked dataset are trained with AlexNet convolutional deep neural network model and the feature packages acquired via the models are processed by the anopheles search algorithm. Subsequently, the efficient features are combined and delivered to random forest (RF) classifier. The proposed approach is implemented in MATLAB. The proposed ADCNN-ASA-RFC provides 91.66%, 69.13%, 34.86%, and 70.13% higher accuracy, 79.13%, 60.33%, and 63.34% higher specificity and 77.13%, 58.45%, 25.86%, and 55.33%, higher sensitivity compared with existing algorithms. At last, the simulation outcomes demonstrate that the proposed system can be able to find the optimal solutions efficiently and accurately with COVID-19 diagnosis.

KEYWORDS

artificial intelligence, computed tomography, COVID-19, medical image and accuracy

1 | INTRODUCTION

The new COVID-19 (2019-nCoV) from Wuhan is presently spreading throughout the world.¹⁻³ Since the documentation of virus in late December 2019,⁴⁻⁵ the number of cases from China that have been imported to other countries has been increasing, and the epidemiological landscape changes daily.⁶⁻⁷ As complete cases of the single SARS-CoV-2 emerge throughout the world,⁸⁻¹³ the entire eye has focused on the seafood market at Wuhan, China, from outbreak source.¹⁴⁻¹⁹ Currently, COVID-19 is the main explanation of death internationally, with the main deaths found in the US, Spain, Italy, China, the United Kingdom, and Iran.²⁰ There are numerous kinds of COVID-19 and SARS-CoV-2 usually found in animals. COVID-19 is exposed on humans, bats, pigs, cat, dog, rodent, and chicken.²¹ The COVID-19 symptoms are pharynges, headache, fever, and so on. Lately, artificial intelligence (AI) has been broadly utilized to accelerate biomedical research.²² By in-depth
learning strategies, AI is employed in numerous applications, like image detection, data classification, and segmentation. X-ray imaging is better than CT, because lesser ionizing radiation, fast data acquisition, accessibility on intensive care units (ICUs), and portability. There is no consensus yet in the combined use of chest CT and CXR in the treatment of COVID-19 pneumonia. The role of machine learning algorithms is to recognize objects. The machine learning algorithm has following advantages, such as it easily identifies trends and patterns, handles multidimensional and multivariety data, does not need human intervention (automation). The random forest classifier is used as the machine learning algorithm. Here, random forest classifier is used to accurately classify the COVID-19 chest X-ray images in to COVID-19, pneumonia, and normal.

1.1 Novelty

The novelty of this article is to detect the type of disease (COVID-19) affected in chest X-ray images and classify the images based on normal, COVID-19, pneumonia with high accuracy less computational time and error rate.

The key contributions of this article are abridged as follows,

- Here, COVID-19 detection and classification framework based on anophoses search optimized AlexNet convolutional deep neural network for random forest classifier is implemented.
- An original dataset is preprocessed with the help of FGLHE technique and fuzzy stacking technique for color enhancement and noise elimination in the input images.
- FGLHE technique and fuzzy stacking technique are combined together and forms into stacked dataset image.
- This stacked dataset are trained with the AlexNet convolutional deep neural network model and therefore the feature packages acquired via the models are processed by the anophes search algorithm.
- Subsequently, the efficient features are combined and delivered to random forest (RF) classifier.
- The proposed approach is implemented on MATLAB.
- The proposed COVID-19 detection and classification framework based on anopheses search optimized AlexNet convolutional deep neural network for random forest classifier performance analysis like accuracy, sensitivity, specificity, precision, and F-score are compared with existing algorithms, like mobile net deep convolutional in social mimic optimization for support vector machine (MNDL-SMO-SVM) and res net deep learning in state of art algorithm for binary classifier (RSDL-STA-BC), computed tomography, and chest X-rays for COVID-19 (HDDNs) Detection, novel process for COVID-19 detect with artificial intelligence under chest X-ray imageries (ResNet) respectively.

The remainder of this article is organized as: Section 2 portrays the recent research works, Section 3 describes the proposed COVID-19 detection and classification framework based on anopheses search optimized AlexNet convolutional deep neural network for random forest classifier, Section 4 demonstrates the experimental results, Section 5 presents the conclusion of this article.

2 LITERATURE REVIEW

Among the recent research works related with COVID-19, a few research works are reviewed here, Jelodar et al., have presented the automated removal of COVID-19 - consideration as social networks and language process scheme supported by topic modeling to discover numerous problems associated with COVID-19 as public opinion. Additionally, the experiments showed that the sample came with an accuracy of 81.15%, which was next precision of many recognized machine learning methods for COVID-19.

Rajaraman et al., have presented the use of repetitive pruned sets of deep learning models to notice the pulmonary expression of COVID-19 with chest X-rays. The knowledge learned was well designed for enhancing performance and generalizations within associated task of classifying chest radiographs as usual, viewing bacterial pneumonia or COVID-19 virus as abnormal. The most effective performance models were pruned again and again for reducing the problem and enhance memory effectiveness.

Wang et al., have suggested a completely unique noise resistant framework to discover noisy labels. They initially advertised a noise robust data loss which was generalization of data loss for segmentation and mean absolute error (MAE) loss. The experimental outcomes demonstrated that: (1) noise resistant matrix loss outstrips existing noise resistant loss functions; (2) COPLE-Net reaches greater execution to next-generation image segmentation networks.
Albahri et al.,\textsuperscript{25} have suggested the artificial intelligence (AI) techniques used on corona virus disease detection and classification in 2019. A comprehensive methodology to evaluate and compare AI processes utilized at entire COVID-19 medical image classification tasks. The experimental outcomes shows that no relevant study has assessed and compared the AI processes utilized in COVID-19 medical image classification tasks.

Abdel-Basset et al.,\textsuperscript{26} have presented to extract quickly from chest X-ray images the similar small regions that have recognizing features of COVID-19. IMPA performance was verified through entry levels among 10 and 100 on nine chest X-ray images compared with five algorithms: Equilibrium Optimizer, whale optimization, sine cosine, Harris-Hawks and Slap Swarm. The outcomes show that the hybrid method outperforms other methods for variety of measurements.

Irfan et al.,\textsuperscript{27} have presented the contribution of hybrid deep neural networks (HDNN), chest radiographs, computed tomography for COVID-19 diagnosis. The aim was to improve the HDNN, computed tomography and X-ray image. It was categorized as three classes: normal, pneumonia, and COVID-19. Firstly, computed tomography along chest X-ray imageries, named “hybrid imageries” (with 1080 \times 1080 resolution) were composed as dissimilar sources, involving Git Hub, the COVID-19 X-ray database, Kaggle, data collection of COVID-19 images and actual med COVID-19 chest X-ray dataset.

Almalki et al.,\textsuperscript{28} have presented an innovative model for the diagnosis of COVID-19 with artificial intelligence on chest X-ray images. The presented algorithm takes various starting residual blocks which was adapted to the information with feature maps of different depths at dissimilar scales. The presented efficient deep learning blocks recycled dissimilar regularization procedures to diminish overfitting based on small COVID-19 data set. Multiscale features were removed at dissimilar levels from presented deep learning model integrated into several machine learning models to authenticate that combination of deep and machine learning models.

Iwendi et al.,\textsuperscript{29} have presented a classification of people with COVID-19 with adaptive neuro-fuzzy inference system. The aimed was to regulate such attributes under early detection of Corona virus disease with adaptive neuro-fuzzy inference system (ANFIS). The presented work calculates the accuracy of different machine learning and chooses the better classifier for COVID-19 identification. The COVID19 dataset was categorized with support vector machine because it accomplished 100% accuracy between the entire classifiers. Additionally, the ANFIS was executed on such classified dataset, resulting at 80% risk prediction for COVID-19.

Iwendi et al.,\textsuperscript{30} have presented a prediction of COVID-19 patient health with powered random forest algorithm. Where, a refined random forest model powered by AdaBoost algorithm was deemed. The presented model utilizes that geographic, travel, health, and demographic data of COVID-19 patient. The model has an accuracy of 94% and F1 score of 0.86 on data set utilized. Data analysis exposes a positive correlation among the gender of patients and deaths denote that most patients are between 20 and 70 years old.

### 2.1 Background of this research work

From the Literature survey,\textsuperscript{22–30} several methods, such as deep learning and machine methods are used to detect and classify the corona virus, it is observed that all the presented works deals with COVID-19 detection and classification with accuracy but faces lot of difficulties, such as noisy in nature, computational time is increased, more expensive to acquire, the assessment and reference process of COVID-19 AI classification occurs multicomplex attribute problem. The histograms of enhanced images spread the full range of gray-scale means there is no saturation and washed-out on output images. But, this proposed method overcomes all this issues and provides more accuracy.

### 3 PROPOSED METHOD FOR AUTOMATIC DETECTION OF COVID-19 WITH CHEST X-RAY IMAGES USING ALEX NET DEEP CONVOLUTIONAL NEURAL NETWORKS AND ANOPHELES SEARCH ALGORITHM IN RANDOM FOREST CLASSIFIER

Here, the proposed method of COVID-19 detection and classification framework using anopheles search optimized AlexNet convolutional deep neural network is implemented for random forest classifier. The fuzzy gray level difference histogram equalization technique and focus stacking technique are used in preprocessing section for enhancing the input image. Then the output of preprocessing are combined together as stacked dataset, which is trained with AlexNet convolutional deep neural network and the trained output into random forest classifier. This random forest classifier classifies as normal, COVID-19 and pneumonia. The overall block diagram for proposed ACDNN-ASA-RFC is shown in Figure 1. From Figure 1, the input COVID-19 chest X-ray images are given to the training and testing phase, then the imageries are preprocessed to eliminate the noises using fuzzy gray level difference histogram equalization technique and then the images are staked using the fuzzy stacking technique. After that the image features are extracted using the AlexNet convolutional deep neural network model and therefore the feature packages acquired via the models are processed by the anopheles search algorithm. Then
FIGURE 1  Block diagram for COVID-19 detection and classification using ACDNN-ASA-RFC for chest X-ray image

The images are classified random forest (RF) classifier. Random forest (RF) classifier accurately classify the images as normal, pneumonia, and normal.

The detail discussion regarding the COVID-19 detection and classification framework using anopheles search optimized AlexNet convolutional deep neural network for random forest classifier is given below in the following section.

3.1  Dataset description

In this article, the three classes of data sets are available for COVID-19 classification such classes are COVID-19 normal, and pneumonia. Here, COVID-19 images are taken from Joseph Paul Cohen database. In which, 70% data set is utilized as training and 30% as testing data in the experimental analysis. This data set consists of total 458 chest images, from this 295 images are detected as COVID-19, 98 images are detected as pneumonia, and 65 images are detected as normal.

3.2  Preprocessing

In this section the preprocessing is employed for separating the redundant content like noise, lights, and background on provided chest X-ray image. The preprocessing step is used for standardizing the input. The dataset is renovated by FGLHE technique and fuzzy stacking procedure.

3.2.1  Fuzzy gray level difference histogram equalization

Gray level differences are blurred to cope with the uncertainties that exist under the input image. A fuzzy gray level clipping range is intended for controlling negligible contrast development. Fuzzy gray level difference histogram balancing techniques are featured to
improve the contrast in the MR clinical image and not to lose their naturalness. In this case, the fuzzy gray level is initially computed for eliminating the uncertainties under the histogram image, and then followed using clipping process to manage the extreme expansion rate.

The size of input image is $m \times n$,

$$l = \{(p,q) \mid 0 \leq p \leq m - 1, 0 \leq q \leq n - 1 \}. \tag{1}$$

The pixel intensity of an input image placed at $(a,b)$ is denoted from $I(a,b) = s$. Here, $s \in \{0, 1, \ldots, L - 1\}$, $L$ implies count of intensity levels.

The following binary space partitioning (BSP) is calculated on $R$ region through $L \times L$ dimension wherever the $L = 5$. The BSP operate from a description of the spatial structure, and it is essential to handle noise intensity in homogeneity and lighting variations on input image. The entire pattern of comparison goes to the $R$ region connected with parameter $T_r = 0.3$.

BSP $P(i_o, i_c)$ described that ratio among intensity of central pixels gray level $i_c$ and neighborhood intensity pixels $i_o$ by the linked coordinates $(m,n)$ on region $R$ is computed as

$$P(i_o, i_c) = \begin{cases} 
1 & \text{if } |i_o - i_c| \leq T_r, \\
0 & \text{otherwise}
\end{cases} \tag{2}$$

Gray level difference among intensity of central pixel $i_c$ through the linked coordinates $(a,b)$ and the intensity of the neighborhood pixel $i_o$ by the linked coordinates $(m,n)$ on region $R$ is computed as

$$X_d(a,b) = \frac{1}{N} \left[ \sum_{m=1}^{L} \sum_{n=1}^{L} P(i(m,n), i(a,b)) \right] \tag{3}$$

where $N$ is represented as the Gaussian field.

The Gaussian membership function uses the described gray level differences vaguely,

$$X_{d}(a,b) = e^{\frac{1}{2} \left( \frac{|X_d(a,b)|}{\sigma} \right)^2} \tag{4}$$

here $\sigma$ implies Gaussian function that may be extended and fitted an exacting weight for assigning the usual interval of gray level differences. This method supports to get that fuzzy gray level differences histogram.

$$h_g(a,b) = \sum_{m=1}^{L} \left( \sum_{n=1}^{L} G_{i_o}(a,b) \delta(i(m,n) - g) \right). \tag{5}$$

g = \{0; 1; 2; 3; \ldots; 255\}.

3.2.2 Stacking techniques

In this section, image stacking mixes multiple images taken or recreates them at changed focal lengths. This is a common way to enhance the standard of photographs within the dataset. This method aims to remove noise as initial image by combining a minimum of two images under same row and separating image.

Let $K_{x,y}(x,F_y,F_z)$ implies marginal Fourier transform of reconstructed target function $\tilde{T}(x,y,z)$ regarding variable $(y,z)$ and then target function on range $x = x_o$ may be improved via two-dimensional inverse Fourier transform of $K_{x,y}(x=x_o,F_y,F_z)$ with respect to $(F_y,F_z)$ which is,

$$\tilde{T}(x=x_o,y,z) = \int_{F_y} \int_{F_z} K_{x,y}(x=x_o,F_y,F_z) \times \exp(jF_y y + jF_z z) dF_y dF_z. \tag{6}$$

here the constant amplitude factor in the Fourier integral is deserted. Comparing Equation (6) find the reconstruction of $K_{x,y}(x=x_o,F_y,F_z)$ is,

$$K_{x,y}(x=x_o,F_y,F_z) = \int_{F_y} S(F_w,F_u,F_v) \times S_w^*(F_w,F_u,F_v) dF_w. \tag{7}$$
where,
\[
F_y = F_u,
\]
\[
F_z = F_v.
\]
(8)

Basically, it can be implemented beyond one-dimensional integral by adding the obtainable discrete values \(F_w\) and Fourier transform digital implementation or their inverse. Thus, the target function is expressed according to the discrete form of reconstruction at \((x = x_n, y, z)\)
\[
\tilde{f}(x = x_n, y, z) = \text{IFFT}\{ \sum_{F_w} \{ \text{FFT}_{uv}\{ S(F_w, u, v) \}\} \cdot S_n(F_w, F_u, F_v) \cdot \Delta F_w \}.
\]
(9)

During this article, MATLAB library is employed to stack system. The initial dataset is stacked on the recreated dataset by Fuzzy system.

### 3.3 AlexNet convolutional neural network model

In this section, the stacked dataset from fuzzy gray level difference histogram equalization technique and fuzzy stacking technique are given to train the AlexNet deep convolutional neural network. AlexNet was proposed by Krizhevsky and trained on 1 million images for classifying images into 1000 different classes; AlexNet is a convolutional deep neural network model with eight layers is deep. Then the five convolutional layers (conv1 to conv5) and completely connected layers are three (fc6 to fc8). The best thing about this model is to execute a successful analysis through decreasing the parameter number, thus decreasing the efficiency of the sample size. The AlexNet model created more successful outcomes, about 50 times less parameters to Squeeze Net model, thus decreasing the model value. However, the knowledge regarding the layer is suggested within the AlexNet convolutional neural network model. The innovative layer has two divisions, specifically the parts of compression and expansion.

The convolutional is the base layer of CNN. In this layer, the input image passes through the filter. The out coming filter values has a characteristic map. This layer uses few cores that slip and finish the pattern to remove the lower and higher level features. The kernel implies \(3 \times 3\) or \(5 \times 5\) shaped matrix that needs to be transferred an input pattern matrix. The convolutional layer output is expressed as:
\[
x^l_j = f \left( \sum_{a=1}^{N} W^{l-1}_{j} \ast Y^{l-1}_{a} + b^l_j \right)
\]
(10)

where \(x^l_j\) implies j feature map on layer l, \(W^{l-1}_{j}\) implies j kernels on layer \(l-1\), \(Y^{l-1}_{a}\) denotes the feature map on layer \(l-1\), \(b^l_j\) implies bias j feature map on layer l, \(N\) indicates total number of features on layer \(l-1\), \((\ast)\) refers vector convolution mode. The second layer after convolutional layer is pooling layer. The pooling layer is frequently used to present maps produced to drop that number of feature maps. The max pooling process chooses only maximal value with the specified array size on every feature map, resulting at decreased output neurons. It is linked with the completely connected layer after the global average pooling layer. The major objective of this layer is to avoid overfitting and network divergence.

The completely connected layer is the significant layer of CNN. This layer acts as multilayered perception. The rectified linear unit (ReLU)\(^{33}\) trigger function is generally utilized in completely connected layer, though the Soft Max trigger function is utilized to forecast output images on final layer of the completely connected layer. The mathematic calculation of two activation functions is below:
\[
\text{ReLU}(x) = \begin{cases} 
0, & x < 0 \\
x, & x \geq 0 
\end{cases}
\]
(11)

\[
\text{Softmax}(x_i) = \frac{e^x_i}{\sum_{j=1}^{m} e^x_j}
\]
(12)

Here, \(x_i\) and \(m\) indicates input data and number of classes, correspondingly. Additionally, it normalizes the dimensional features of the images with the use of computing variance as well as mean value of the input image samples. In this, two parameters \(\beta, \varphi\) are introduced for recovering the feature distribution function (\(\tilde{f}\)) using parameter training. The calculation formula is mentioned as Equation (2) and Equation (3),
\[
p_i = \varphi \tilde{t}_i + \beta,
\]
(13)

\[
\tilde{t}_i = \frac{t_i - \mu}{\sqrt{\sigma^2 + Y}}
\]
(14)
where $\mu, \sigma$ denotes the sample mean value and variance value. $\Upsilon$ Denotes the constant value and $i$ represents the total number of samples. Also, $\beta, \varphi$ represent the parameter used for feature extraction. By using AlexNet various COVID-19 features are extracted. While detecting the images some errors may occurs that will reduce the accuracy. For reducing the error rate, the parameters of the AlexNet is optimized using the Anopheles search algorithm.

Anopheles search algorithm is a new meta-heuristic algorithm for optimizing the errors in the networks. Here anopheles search algorithm is used to optimize the parameters of the AlexNet such as $\beta, \varphi$ for improving accuracy and to lessen that error rate during classification process. In this $\beta$ is recycled to upsurge the accuracy and $\varphi$ is used to lessen that error rate. Anopheles search algorithm (ASA) is a process evolved in every problem creates a solution with AS algorithm that moves to the optimal solution. Every solution decides the variance on optimal value by likening the local value acquired on final repetition.

3.3.1 Step by step procedure of ASA

Here, optimizations are takes place based on the anopheles mosquitoes that are close to three tents. Some anopheles mosquitoes are positioned that not carriers near three tents. At initial tent, a malaria-infected child with parasitism gets ready to transmission. At second, a malaria-infected child with parasitism not ready to transmission. At third, a healthy child is existed. The mosquito sniffs the smell of children’s body through the air. The outcomes demonstrate that the mosquitoes have inspired to the smell of the boy who is in the first tent twice as much as other children. By this process, the Anopheles is used to optimize the parameters of the AlexNet for increasing accuracy and to reduce the error rate. Figure 2 portrays the flow chart of ASA-AlexNet.

**Step 1: Initialization.**

The initial anopheles is initialized for optimizing the parameters of the AlexNet. When the term $P$ in ASA expresses the population, the parameter $B$ articulates the better optimal value, and parameter $I$ implies number of iterations. Decision variables consist of lower bound $lb$ and upper bound $ub$ value. The following equations are used to implement the AS algorithm.

\[ P_i = (1, j) = lb_j + x \cdot (ub_j - lb_j) ; j = 1, 2, \ldots, N \] \hspace{1cm} (15)

\[ P_i = P_i + x \cdot (P_i) \] \hspace{1cm} (16)

In Equation (11), an inverse distance is taken among $X_i$ and point $Y_i$, then $C$ represents optimal point on solution space according to the location $X_i$ may be consequent as below,

\[ O_{X_i Y_i} = \frac{1}{\text{Dist}(X_i Y_i)} \times \log(fitness(Y_i)) + b, 0 \leq b \leq 0.5. \] \hspace{1cm} (17)

**Step 2: Random generation.**

Mosquitoes are distributed randomly under search space. Then the distance calculation equation is given in Equation (18)

\[ \text{Dist}(X_i Y_i) = \sqrt{\sum_{j=1}^{n} (X_j - Y_j)^2}. \] \hspace{1cm} (18)

If $b = 0.5$, the $O$ will enlarge. The closer value of $b$ to their upper one of the Anopheles movement maximized. If $b = 0$, then the distance and optimality are employed for computing anopheles movement.

**Step 3: Fitness function.**

This is used to attain that objective function for increasing accuracy and lessen the error rate. The fitness function is expressed in Equation (19),

\[ \text{Fitness function}(X_i Y_i) = \max_{\beta}(\text{Accuracy}), \min_{\varphi}(\text{Error rate}), \] \hspace{1cm} (19)

where $\beta$ is represented as the accuracy and $\varphi$ is represented as the error rate.

**Step 4: Updation of ASA for increasing accuracy and decreasing error rate.**

In this optimization, the parameters are determined by obtaining the perceived odor density of each mosquito and moved toward the optimum point depending on the accomplished value. Then the accuracy is increased by reducing the error rate and the maximization equation is given in Equation (20)
\[ \text{Maximize}(\text{Accuracy}) (X_i), \quad X_i \in x_i, i = 1, 2, \ldots N. \] (20)

where \( \text{Maximize}(\text{Accuracy}) (X_i) \) is represented as the objective function, \( X_i \) refers set of decision variables, \( x_i \), and \( i \) refers set of feasible values of every decision variable, \( N \) refers number of decision variables. If decision variables have continual, then it is demarcated with upper \( (U_{x_i}) \) including lower bounds \( (L_{x_i}) \), which is given in Equation (21), (22),

\[ L_{x_i} \leq x_i \leq U_{x_i}. \] (21)

\[ x_i \in \{x_1, x_2, \ldots, x_n\}. \] (22)

Then the error rate minimized equation is given in Equation (23)

\[ \text{Minimize}(\text{Error rate}) (Y_i) = x_1^2 + (x_2 - 1)^2. \] (23)

Equations (18) and (21) are known as the optimization equation for achieving the objective function. The parameters are optimized based on the odors of anopheles.
**Step 5**: Termination.

Finally, the objective function is utilized for maximizing the accuracy by minimizing the computational time with fault. At last, the best values are selected from the AlexNet classifier through ASA mechanism, which are effectively classified the input chest X-ray imageries as COVID-19 and pneumonia images.

### 3.4 Classification using random forest classifier

In this section, random forest classifier is employed to the COVID-19 classification. The random forest classifier is a stochastic search system to solve optimal solutions in huge and complex search spaces. It is effectively recycled for a popular evolutionary algorithm (EA) classification. Every individual is an input data encoding named chromosomes. The search for a better solution is guided through an objective function named fitness function. High fitness function designated solutions have more capacity to create innovative solutions to lower fitness value ones, while weak fitness function ones will be phased out.

\[
F = \frac{C_i}{T_i},
\]

where \( F \) is the fitness function, \( C_i \) indicates the number of classified instance and \( T_i \) the training samples.

The random forest classifier (RFC) is most effective cooperative learning procedure proven to the most popular and powerful systems under model recognition and machine learning for categorizing high dimensions and skewed issues.

Assume the learning set

\[
L = ((N_1, M_1), \ldots, (N_m, M_m)).
\]

Consider \( m \) is the vectors, \( N \in X \) here \( X \) is set of numeric or symbolic observations and \( M \in Y \) where \( Y \) refers class label set. A classifier denotes mapping \( X \rightarrow Y \). A novel input vector is categorized by every individual tree of random forest.

### 4 EXPERIMENTAL RESULTS AND DISCUSSION

This section describes about simulation result of proposed system on unique chest imageries of COVID-19 virus infection as normal chest images and pneumonia. The simulations are performed on a PC through the Intel Core i5, 2.50 GHz CPU, 8GB RAM and Windows 7. The proposed system is carried out in MATLAB. Here, evaluation metrics, like accuracy, sensitivity, specificity, and precision F-score are analyzed. The performance analyses are likened with existing mobile net deep convolutional in social mimic optimization for support vector machine (MNDL-SMO-SVM) and res net deep learning in state of art algorithm for binary classifier (RSDL-STA-BC), computed tomography, chest X-rays for COVID-19 (HDDNs) detection, new approach for COVID-19 diagnosis with artificial intelligence on chest X-ray imageries (ResNet) algorithms respectively.

#### 4.1 Dataset description

In this article, the three classes of data sets are available for COVID-19 classification such classes are COVID-19 normal, and pneumonia. In this the COVID-19 dataset is taken from the Joseph Paul Cohen database. In the experimental investigation, 70% dataset is recycled as training, 30% is recycled as testing data. This data set consists of total 458 chest images, from this 295 images are detected as COVID-19, 98 images are detected as pneumonia, and 65 images are detected as normal.

The simulation parameters of proposed method are verified on Table 1.

| Parameter         | Value     |
|-------------------|-----------|
| Software used     | MATLAB    |
| Population size   | 20        |
| Iteration parameter | 10       |
| Global parameter  | 1000      |
| Image size        | 224 × 224 |
| Learning rate     | \(10^{-5}\) |
4.2 | Performance metrics

4.2.1 | Accuracy

The measure of the overall effectiveness of classification system is named as accuracy

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}.
\]

(26)

4.2.2 | Specificity

For identifying patterns of a negative class is computed to measure the ability of the classifier

\[
\text{Specificity} = \frac{TN}{TN + FP}.
\]

(27)

4.2.3 | Sensitivity

For identifying patterns of a positive class is computed to measure the ability of the classifier

\[
\text{Specificity} = \frac{TN}{TN + FN}.
\]

(28)

4.2.4 | F-measure

F-measure creates the use of recall and precision is described as

\[
F_{\text{measure}} = 2 \times \frac{\text{recall} \times \text{precision}}{\text{recall} + \text{precision}}.
\]

(29)

Recall = \[
\frac{TP}{TP + FN}.
\]

(30)

Precision = \[
\frac{TP}{TP + FP}.
\]

(31)

4.2.5 | Positive predictive value

\[
\text{PPV} = \frac{TP}{TP + TN}.
\]

(32)

4.2.6 | Negative predictive value

\[
\text{NPV} = \frac{TN}{TN + FN}.
\]

(33)

Although specimens belonging to the selected class are properly recognized via classifier, such specimens are placed in TP codes. Other models belonging to correctly recognized opposite classes are within the TN codes within the confusing matrix.

Figure 3 shows the separate chest imageries of COVID-19 as normal chest imageries and pneumonia. Here, the input COVID-19 chest X-ray images are given to the training and testing phase, then the images are preprocessed to eradicate the noises using fuzzy gray level difference histogram equalization technique and then the images are staked using the fuzzy stacking technique. After that the image features are extracted using
the AlexNet convolutional deep neural network model and therefore the feature packages acquired via the models are processed by the anoph- 
les search algorithm. Then the images are classified random forest (RF) classifier. Random forest (RF) classifier accurately classifies the images as normal, pneumonia, and normal.

Table 2 shows the performance metrics of the Accuracy through original data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 20.25%, 35.86%, 28.42%, and 15.85% higher accuracy than existing methods respectively. For normal, the proposed ADCNN-ASA-RFC provides 5.58%, 20%, 23.97%, 45.97% higher accuracy than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 12.19%, 2.22%, 35.97%, 37.55% higher accuracy than existing methods respectively.

Table 3 shows the performance metrics of accuracy through stacked data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method. For COVID-19, the proposed ADCNN-ASA-RFC provides 25.06%, 32.08%, 42.54%, 34.87% better accuracy over existing models. For normal, the proposed ADCNN-ASA-RFC provides 36.97%, 27.06%, 38.97%, 29.97% better accuracy over existing models. For pneumonia, the proposed ADCNN-ASA-RFC provides 34.06%, 26.08%, 37.97%, 29.97% better accuracy over existing models.

Table 4 shows the performance metrics of precision with original data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing methods, such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet, respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 16.07%, 33.09%, 27.65%, and 43.54% higher precision than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 13.65%, 25.97%, 31.05%, and 28.97% higher precision than existing such as existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 25.73%, 24.65%, 38.96%, and 32.91% higher precision than existing methods.

**Table 2** Performance Analysis of accuracy using original data

| Diseases     | MNDL-SMO-SVM | RNDL-STA-BC | HDNNs | ResNet | ADCNN-ASA-RFC (proposed) |
|--------------|--------------|-------------|-------|--------|--------------------------|
| COVID-19     | 75           | 69          | 74    | 76     | 96                       |
| Pneumonia    | 66           | 73          | 65    | 84     | 97                       |
| Normal       | 76           | 59          | 75    | 64     | 95                       |

**Table 3** Performance analysis of accuracy using stacked data

| Diseases     | MNDL-SMO-SVM | RNDL-STA-BC | HDNNs | ResNet | ADCNN-ASA-RFC (proposed) |
|--------------|--------------|-------------|-------|--------|--------------------------|
| COVID-19     | 67           | 57          | 59    | 76     | 98                       |
| Pneumonia    | 59           | 64          | 73    | 59     | 97                       |
| Normal       | 63           | 73          | 65    | 74     | 96                       |

**FIGURE 3** Proposed method for convolutional deep neural network model
Table 5 shows the performance metrics of the precision with stacked data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet, respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 33.05%, 26.05%, 32.08%, and 36.97% higher precision than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 34.07%, 26.65%, 31.86%, and 32.97% higher precision than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 26.65%, 34.08%, 38.96%, and 22.75% higher precision than existing methods.

From Table 6 shows that the performance metrics of the specificity using original data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 34.8%, 26.07%, 45.86%, and 32.97% higher specificity than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 36.76%, 26.55%, 38.97%, and 32.08% higher specificity than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 26.65%, 34.08%, 38.96%, and 22.75% higher specificity than existing methods.

Table 7 shows the performance metrics of the specificity using stacked techniques for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 20.76%, 23.65%, 26.97%, and 20.94% higher specificity than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 24.97%, 35.86%, 32.97%, and 21.97% higher specificity than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 24.97%, 34.87%, 55.35%, 32.97%, and 25.97% higher specificity than existing methods.

Table 8 shows the performance metrics of the sensitivity using original data for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet, respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 25.64%, 19.51%, 34.97%, and 56.63% higher sensitivity than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 37.03%, 5.12%, 35.34%, and 27.08% higher sensitivity than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 3.79%, 46.42%, 24.64%, and 26.08% higher sensitivity than existing methods.

Table 9 shows the performance metrics of the sensitivity using stacked techniques for detecting and classifying of input dataset image as COVID-19, normal, pneumonia, and proposed system is comparing with various existing method such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, ResNet, respectively. For COVID-19, the proposed ADCNN-ASA-RFC provides 25.31%, 43.97%, 46.97%, and 6.94% higher sensitivity than existing methods. For normal, the proposed ADCNN-ASA-RFC provides 27.98%, 43.97%, 27.83%, and 23.08% higher sensitivity than existing methods. For pneumonia, the proposed ADCNN-ASA-RFC provides 26.97%, 34.09%, 28.08%, and 44.04% higher sensitivity than existing methods.

**TABLE 4** Performance analysis of precision using original data

| Diseases   | MNDL-SMO-SVM | RNDL-STA-BC | HDNNs | ResNet | ADCNN-ASA-RFC (Proposed) |
|------------|--------------|-------------|-------|--------|--------------------------|
| COVID-19   | 66           | 58          | 62    | 63     | 91                       |
| Pneumonia  | 57           | 59          | 65    | 59     | 92                       |
| Normal     | 53           | 67          | 57    | 60     | 94                       |

**TABLE 5** Performance analysis of precision using stacked data

| Diseases   | MNDL-SMO-SVM | RNDL-STA-BC | HDNNs | ResNet | ADCNN-ASA-RFC (Proposed) |
|------------|--------------|-------------|-------|--------|--------------------------|
| COVID-19   | 57           | 67          | 59    | 54     | 92                       |
| Pneumonia  | 49           | 55          | 72    | 67     | 90                       |
| Normal     | 74           | 62          | 66    | 63     | 93                       |

**TABLE 6** Performance Analysis of specificity using original data

| Diseases   | MNDL-SMO-SVM | RNDL-STA-BC | HDNNs | ResNet | ADCNN-ASA-RFC (Proposed) |
|------------|--------------|-------------|-------|--------|--------------------------|
| COVID-19   | 57           | 78          | 58    | 63     | 89                       |
| Pneumonia  | 76           | 54          | 64    | 73     | 86                       |
| Normal     | 47           | 51          | 65    | 54     | 90                       |
Table 7 shows the performance analysis of specificity using stacked techniques for diseases COVID-19, pneumonia, and normal. The proposed ADCNN-ASA-RFC provides higher specificity than existing methods such as MNDL-SMO-SVM, RNDL-STA-BC, and HDNNs, respectively.

Table 8 shows the performance analysis of sensitivity using original data for the same diseases. The proposed ADCNN-ASA-RFC provides higher sensitivity than existing methods such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, and ResNet, respectively.

Table 9 shows the performance analysis of sensitivity using stacked techniques for the same diseases. The proposed ADCNN-ASA-RFC provides higher sensitivity than existing methods such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, and ResNet, respectively.

Table 10 shows the performance analysis of F-score using original data for detecting and classifying input dataset image as COVID-19, normal, pneumonia, and proposed system compared with various existing methods such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, and ResNet, respectively. The proposed ADCNN-ASA-RFC provides higher F-score than existing methods for COVID-19, normal, and pneumonia.

Table 11 shows the performance analysis of F-score using stacked techniques for the same diseases. The proposed ADCNN-ASA-RFC provides higher F-score than existing methods such as MNDL-SMO-SVM, RNDL-STA-BC, HDNNs, and ResNet, respectively.
CONCLUSION

In this article, COVID-19 detection and classification framework based on anopheles search optimized AlexNet convolutional deep neural network for random forest classifier is successfully implemented. The proposed work is simulated in MATLAB. The proposed ADCNN-ASA-RFC provides higher precision 86.66%, 66.33%, and 70.66% and higher F-score 80.33%, 53.33%, and 59.66% compared with existing algorithms like mobile net deep convolutional in social mimic optimization for support vector machine (MNDL-SMO-SVM) and res net deep learning in state of art algorithm for binary classifier (RSDL-STA-BC). The performance of proposed model is evaluated through expert radiologists and it is prepared through a greater database. This model is recycled on remote locations in countries exaggerated by COVID-19 to overwhelm the shortage of radiologists. In this way, this method is applicable in the clinical use.

This ADCNN-ASA-RFC is appropriate in the applications of real time. The ADCNN-ASA-RFC is also able to classify the patients based on COVID-19 by collecting real time data from various health care services, such as medical diagnose center, hospital. Not only that, this method is utilized to proficiently classify the input dataset image as COVID-19, normal, pneumonia. Furthermore, this method assesses the accuracy of various deep learning and machine strategies employed to detect the COVID-19 at beginning stage. This method proves better prediction along classification tasks when compared with other classifier methods. These tasks are applied in COVID-19 for disease probability prediction, screening, diagnosis, treatment guidance, and complication management. Several methods have been applied to fulfill these clinical purposes using their own specific capabilities. Hence, this application represents a promising tool to aid the stratification of COVID-19 patients.

Future work will focus on generating a pipeline that connects chest X-ray scanning computer vision models through such sorts of healthcare and demographic data process models. These models will be incorporated to requests that help the evolution of mobile healthcare. It can deliver a step toward semiautonomous diagnostic system and give rapid detection for regions exaggerated by COVID-19.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Sumaiya Begum Akbar https://orcid.org/0000-0002-0931-5709
Kalaiselvi Thanupillai https://orcid.org/0000-0002-9423-0223

REFERENCES

1. Rothe C, Schunk M, Sothmann P, et al. Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. N Engl J Med. 2020;382(10):970-971.
2. Gostic K, Gomez AC, Mummah RO, Kucharski AJ, Lloyd-Smith JO. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. Elife. 2020;9:e55570.
3. Khan N, Fahad S, Naushad M, Faisal S. Critical review of COVID-19 in the world. SSRN Electron J. 2020.
4. Cohen J. Wuhan seafood market may not be source of novel virus spreading globally. Science. 2020;10(1.11.26).
5. Yi Y, Laginot P, Ye S, Li E, Xu R. COVID-19: what has been learned and to be learned about the novel coronavirus disease. Int J Biol Sci. 2020;16(10):1753-1766.
6. Rajagopalan P. The deadly Corona virus (COVID-19). J Commun Dis. 2020;52(01):78-81.
7. Keni R, Alexander A, Nayak P, Mudgal J, Nandakumar K. COVID-19: emergence, spread, possible treatments, and global burden. Front Public Health. 2020;2:16.
8. Jin Y, Yang H, Ji W, et al. Virology, epidemiology, pathogenesis, and control of COVID-19. Viruses. 2020;12(4):372.
9. Ng KT, Takebe Y, Kamarulzaman A, Tee KK. Genome sequencing and phylogenetic reconstruction reveal a potential fourth rhinovirus species and its worldwide distribution. Arch Virol. 2021;166(1):225-229.
10. Mythili S, Thiagarajah K, Rajesh P, Shajin FH. Ideal position and size selection of unified power flow controllers (UPFCs) to upgrade the dynamic stability of systems: an antlion optimiser and invasive weed optimisation algorithm. HKIE Trans. 2020;27(1):25-37.
11. Rajesh P, Shajin F. A multi-objective hybrid algorithm for planning electrical distribution system. Eur J Electr Eng. 2020;22(4-5):224-509.
12. Shajin FH, Rajesh P. Trusted secure geographic routing protocol: outsider attack detection in mobile ad hoc networks by adopting trusted secure geographic routing protocol. Int J Pervas Comput Commun. 2020.
13. Thota MK, Shajin FH, Rajesh P. Survey on software defect prediction techniques. Int J Appl Sci Eng. 2020;17(4):331-344.
14. Fuchs A, Lanzi D, Beilstein CM, et al. Clinical recommendations for in-hospital airway management during aerosol-transmitting procedures in the setting of a viral pandemic. Best Pract Res Clin Anaesthesiol. 2021;35(3):333-349.
15. Mostafiz R, Uddin MS, Alam NA, Reza MM, Rahman MM. COVID-19 detection in chest X-ray through random forest classifier using a hybridization of deep CNN and DWT optimized features. J King Saud Univ Comput Inf Sci. 2020.
16. Korinek A, Stiglitz JE. COVID-19 driven advances in automation and artificial intelligence risk exacerbating economic inequality. BMJ. 2021;372:n367.
17. Leslie D, Mazumder A, Pepin A, Wolters MK, Hagerty A. Does “AI” stand for increasing inequality in the era of COVID-19 healthcare? BMJ. 2021;372:n304.
18. Cave S, Whittlestone J, Nyrup R, Calvo RA. Using AI ethically to tackle COVID-19. BMJ. 2021;372:n364.
19. Wymants L, Van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of COVID-19: systematic review and critical appraisal. BMJ. 2020;369.
20. Kalaiselvi T, Rahimunnisa K, Sumaiya Begum A. Pulse coupled neural networks for image processing – A review. Int J Adv Sci Technol (IJAST). 2020;29(7s):4411-4416.
21. Begum AS, Savitha S, Shahila S, Sharmila S. Diagnosis of leaf disease using enhanced convolutional neural network. Int J Innovat Res Appl Sci Eng. 2020;3(12):579-586.
22. Jeelodar H, Wang Y, Orji R, Huang S. Deep sentiment classification and topic discovery on novel coronavirus or COVID-19 online discussions: NLP using LSTM recurrent neural network approach. IEEE J Biomed Health Inform. 2020;24(10):2733-2742.
23. Rajaraman S, Siegelman J, Alderson P, Folio L, Folio L, Antani S. Iteratively pruned deep learning ensembles for COVID-19 detection in chest X-rays. IEEE Access. 2020;8:115041-115050.
24. Wang G, Liu X, Li C, et al. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images. IEEE Trans Med Imaging. 2020;39(8):2653-2663.
25. Albahri OS, Zaidan AA, Albahri AS, et al. Systematic review of artificial intelligence techniques in the detection and classification of COVID-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects. J Infect Public Health. 2020;13(10):1381–1396.
26. Abdel-Basset M, Mohamed R, Elhoseny M, Chakrabortty R, Ryan M. A hybrid COVID-19 detection model using an improved marine predators algorithm and a ranking-based diversity reduction strategy. IEEE Access. 2020;8:79521-79540.
27. Irfan M, Iftikhar MA, Yasin S, et al. Role of hybrid deep neural networks (HDNNs), computed tomography, and chest X-rays for the detection of COVID-19. Int J Environ Res Public Health. 2021;18(6):3056.
28. Almalki YE, Qayyum A, Irfan M, et al. A novel method for COVID-19 diagnosis using artificial intelligence in chest X-ray images. In Healthcare. Multidisciplinary Digital Publishing Institute:2021;9(5):22.
29. Iwendi C, Mahboob K, Khalid Z, Javed AR, Rizwan M, Ghosh U. Classification of COVID-19 individuals using adaptive neuro-fuzzy inference system. Multimed Syst. 2021;1–15.
30. Iwendi C, Bashir AK, Peshkar A, et al. COVID-19 patient health prediction using boosted random forest algorithm. Front Public Health. 2020;8:357.
31. Kandhway P, Bhandari AK, Singh A. A novel reformed histogram equalization based medical image contrast enhancement using krill herd optimization. Biomed Signal Process Control. 2020;56:101677.
32. Li W, Wang G, Chen X, Yin X, Hu X, Blurring-effect-free CNN for optimization of structural edges in focus stacking. Proceedings of the 2019 IEEE International Conference on Image Processing (ICIP): 2019;4634-4638; IEEE.
33. Baloochian H, Ghaffary HR, Balochian S. Metaheuristic anopheles search algorithm. Evol Intel. 2020;13(3):511-523.
34. Türküç M, Ergen B, Cömert Z. COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. Comput Biol Med. 2020;121:103805.
35. Jain G, Mittal D, Thakur D, Mittal MK. A deep learning approach to detect COVID-19 coronavirus with X-ray images. Biocy Biomed Eng. 2020;40(4):1391-1405.

How to cite this article: Akbar SB, Thanupillai K, Sundararaj S. Combining the advantages of AlexNet convolutional deep neural network optimized with anopheles search algorithm based feature extraction and random forest classifier for COVID-19 classification. Concurrency Computat Pract Exper. 2022;34(15):e6958. doi: 10.1002/cpe.6958