Check for updates

A cad system design based on HybridMultiscale convolutional Mantaray network for pneumonia diagnosis

Durga Prasad Mannepalli[1] · Varsha Namdeo[2]

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

Abstract

Pneumonia is one of the diseases that people may encounter in any period of their lives. Recently, researches and developers all around the world are focussing on deep learning and image processing strategies to quicken the pneumonia diagnosis as those strategies are capable of processing numerous X-ray and computed tomography (CT) images. Clinicians need more time and appropriate experiences for making a diagnosis. Hence, a precise, reckless, and less expensive tool to detect pneumonia is necessary. Thus, this research focuses on classifying the pneumonia chest X-ray images by proposing a very efficient stacked approach to improve the image quality and hybridmultiscale convolutional mantaray feature extraction network model with high accuracy. The input dataset is restructured with the sake of a hybrid fuzzy colored and stacking approach. Then the deep feature extraction stage is processed with the aid of stacking dataset by hybrid multiscale feature extraction unit to extract multiple features. Also, the features and network size are diminished by the self-attention module (SAM) based convolutional neural network (CNN). In addition to this, the error in the proposed network model will get reduced with the aid of adaptivemantaray foraging optimization (AMRFO) approach. Finally, the support vector regression (SVR) is suggested to classify the presence of pneumonia. The proposed module has been compared with existing technique to prove the overall efficiency of the system. The huge collection of chest X-ray images from the kaggle dataset was emphasized to validate the proposed work. The experimental results reveal an outstanding performance of accuracy (97%), precision (95%) and f-score (96%) progressively.

✉ Durga Prasad Mannepalli
   dp.mannepalli@gmail.com

1   Research Scholar, Department of Computer Science & Engineering, Sarvepalli Radhakrishnan University, Bhopal, Madhya Pradesh, India

2   Department of Computer Science & Engineering, Sarvepalli Radhakrishnan University, Bhopal, Madhya Pradesh, India

1 Introduction

The lungs of any human can be affected by an acute respiratory infection called Pneumonia
[27]. This pneumonia can be categorised as bacterial pneumonia and viral pneumonia. In
general, bacterial pneumonia holds several acute symptoms. The bacterial and viral pneumonia
are significantly differentiated by the way of treatment. The bacterial pneumonia can be treated
with antibiotic therapy. Viral pneumonia can be recovered without any others help. [9] The
hazardous of pneumonia could be very severe and needs time for regaining the health of
patient. Hence, crucial research is required for developing a novel computer-aided diagnosis
(CAD) approach that will lessen mortality associated with pneumonia, specifically child
mortality, in the emerging biosphere.

Clinical experts use various images such as magnetic resonance imaging (MRI), chest X-
ray, and CT for diagnosing pneumonia. A chest X-ray plays an essential role in the pneumonia
diagnosis system due to the lesser cost and easier access. Furthermore, all the hospitals have
radiology imaging systems hence the appropriateness and accessibility of the radiography-
based methods are very high [5, 11, 14]. Radiology experts need to put more effort and time
into analysing the X-ray image. Hence, different computer algorithms have developed for
analysing the x-ray images. In addition, different CAD designs have industrialized for
providing an intuition of x-ray images. But doctors do not obtain adequate data from these
tools to make a decision [13].

The CAD systems are usually developed with certain statistical analysis, machine learning
and deep learning tools. These approaches are employed for tackling multifaceted computer
vision issues in the field of medical imaging such as lungs segmentation [4, 10, 25], lungs
classification and so on [3, 6, 20]. Deep learning is considered as a significant division in the
arena of artificial intelligence, which obtains raw data as input and represents the features using
the layered concepts. At last, it maps the input towards the target class based on the learned
features. These deep learning approaches do not necessitate human interpretation [16]. The
primary stages that are included in CAD systems are pre-processing, regions of interest (ROI)
extraction, feature extraction and classification. The medical image analysis utilized deep
learning approaches because of their suitability in the replacement process of feature extraction
and classification ability of chest X-ray than that of traditional CADs [15]. CNN is conceded as
the best image classification model because it contains few parameters and interconnections.
Hence, the training process of CNN is much easier than that of remaining neural networks.
Here, the high-level features are learned with deeper structures for ensuring the best classifi-
cation accuracy [21, 29].

The scarcity of X-rays makes the training process of deep neural networks a more
challenging task. Thus, most of the recent pneumonia detection approaches adopt transfer
learning as a feasible solution in this condition. Some of the transfer learning models are
Xception, VGG16, VGG19, ResNet and its variants, Inception-v3 and its variants, MobileNet,
DenseNet, and NASNet [12].

Many researchers are interested in CNN, and this application has a tremendous outcome in
a medical image classification task. The deep learning models need huge data variability to
provide a reasonable presentation. Also, the compact CNN networks do not extract adequate

features due to their shallow structure. Hence, how to increase the capability of such model in classification task is still a big challenge. The CNN module can clear up the pattern recognition issues and object extraction problems and also this network is also inclined with vast processing capacity. In present, all around the world is affected due to the pandemic of COVID19 that is family of pneumonia. Many of the clinics employ such kinds of clinical data for managing their patient or healthcare data. This usually makes large amounts of data that take text, numbers, charts and images. Inappropriately, this information will get used infrequently for supporting decision-making in the hospital. To pinpoint the classification, it grasps plenty of times. Therefore there is a demand for a classification strategy with prohibitive accuracy. The above mentioned has motivated us to propose a hybrid effective, optimized hybrid neural network model for the pneumonia classification and it seems to be a major motivation of this research. Therefore, a robust and precise model is needed to extract features and the classification of Pneumonia for the rapid diagnosis.

In this study, we have embedded the self-attention based convolution mantaray network model due to its speedy process of training. It recommended notably few training parameters when measuring up with existing techniques and this method is computationally efficient. The multiscale feature extraction unit is emphasised in the first convolutional layer of the CNN for the extraction of deep feature and the optimization algorithm is jointly contributed in the self attention based CNN for the optimal selection of hyper parameters or the loss function reduction. CNN is emphasised as an artificial neural network (ANN) and it is augmented to undertake the images automatically. The CNN network model is integrated with multiple layers. Sharing weights and connectivity among the neurons is the essential role of CNN networks. The concepts of connectivity in neural network diminish the complexity of the network. In a real-world application, the deep learning schedule had procured massive success in classification methods. The several hindrances may moderate its usage in certain applications. It requires an enormous amount of labelled data to train the parameters. The major contribution of this research is to upsurge the accuracy of pneumonia detection model by selecting the more discriminant deep features with less computational time. The major contributions of the proposed work include the following:

- An efficient hybrid dataset restructuring strategy is intended to improve image quality and minimize the interference and dataset reconstruction. The stacked dataset is further utilised for training and this will leads to extract the efficient features.
- By the intrusion of a multiscale feature extraction unit in the first convolutional layer, the narrow network model is presented to extract more discriminant deep features with a smaller number of layers and diminish the extracted features' dimension.
- The self attention mechanism is intended in the feature extraction model to capture the long dependences of deep features, reduce the network depth, and improve the classification accuracy.
- To improve the classification accuracy by selecting optimal hyper parameters through meta-heuristic approaches.

The remaining sections of the paper are organized as follows: Sections 2 provides the survey of recently published papers relevant to pneumonia classification. Section 3 describes the proposed methodologies with three stages. Section 4 validates the simulation part along with results and discussion. Section 5 affords the conclusion and future scopes of the proposed article.

2 Literature survey

2.1 Some of the recent related works are discussed as follows

Rahman et al. [22] detected bacterial and viral pneumonia by introducing deep CNN with transfer learning. Initially, the chest X-ray images were pre-processed to perform normalization based on the morals of pre-trained models. These pre-processed images were utilized to train four CNNs including AlexNet, ResNet18, DenseNet201, and SqueezeNet, based on transfer learning. Data augmentation (rotation, scaling and translation) had intensified to amplify the classification accuracy. The work had been validate with chest X-ray dataset and the DenseNet201 with normal, bacteria and viral pneumonia attains an accuracy, sensitivity, precision and f score of 93.3%, 93.2%, 93.7%, 93.5%, ResNet18 of 87.7%, 88%, 87.5%, 90%,AlexNet of 88.4%, 83%, 88%, 88% and Squeeze Net of 86%, 85%, 87%,86% respectively. Sirazitdinov et al. [24] investigated the prospects of machine learning solutions in automatic pneumonia detection and localization application. Furthermore, they enriched two CNNs such as Mask R-CNN and RetinaNet to efficiently detect and localize pneumonia. Here, RetinaNet and Mask R-CNN have been utilized as primary and secondary model respectively to adjust the portions of pneumonia. The network parameters of the proposed model had adjusted using a non-maximum suppression (NMS) approach. By the experimental validation, the developed methods attain a precision, recall and f score of 75%, 79% and 77%, respectively.

Zhang et al. [30] formulated anomaly detection as a one-class classification task that differentiates viral pneumonia from non-viral one. To do that, a confidence-aware anomaly detection unit (CAAD) has been proposed. When that unit produced smaller confidence score then the input would be categorized as viral pneumonia. In that approach, the whole viral pneumonia classes were treated as anomalies to evade specific modelling for each viral pneumonia categories. From that study, the performance such as area under the receiver operator curves (AUC), sensitivity, specificity, and, accuracy had validated. From the evaluation the developed model attains an AUC of 83.61%, accuracy of 72.77%, sensitivity of 71.70% and specificity of 73.83%, respectively.

Yee and Raymond [28] suggested a CNN network for the pneumonia diagnosis using chest X-ray images. Here, the author had introduced three models, such as K-nearest neighbour (KNN), neural network (NN) and support vector machines (SVM) consecutively. The performance of KNN, NN and SVM attains an accuracy of (83%), (84%) and (84%) also the three model had procure a precision of (83%), (74%) and (92%), recall of (84%), (75%) and (83%), f score of (84%), (75%) and (90%) respectively. Ying et al. [26] had developed deep learning techniques for the efficient diagnosis of COVID-19 with the rid of CT images. The process of CT diagnosis had progressed by three essential stages. Initially, the extraction of main regions of lungs and then details relation extraction neural network (DRE-net) was developed to extract top-K details assimilated in the CT images. Finally, the prediction of image level was coordinated to attain a patient level. From the experimental analysis it had been seen that VGG16, Dense Net and ResNet had been compared with DRE-Net and it attains an accuracy of (91%), (87%), (90%), (95%), precision of (80%), (76%), (81%), (79%), recall of (89%), (93%), (93%), (96%) and f score of (84%), (83%), (86%), (87%) respectively.

Mahmud et al. [17] proposed deep CNN to automatically detect COVID-19 and different kinds of pneumonia automatically using smaller chest X-ray datasets. That proposed model utilized different dilation rates in the depth-wise convolution to efficiently extract expanded

features of chest X-rays. Here, the early training stage had learned using certain added fine-tuning layers and they were used less amount of chest X-rays for further training. In initial, fine-tuning layers can be transferred with the aid of training phase and they were further trained to diagnose pneumonia. That proposed prediction model has been optimized using a stacking algorithm. At last, gradient-based discriminative localization had used for distinguishing the anomalous portion of X-ray images that reflected various kinds of pneumonia. The developed stacked multi-resolution CovXNet method attains normal images with 97%, viral with 87.3%, bacterial with 94%, both viral and bacterial with 89%, normal, viral and bacterial with 90%, respectively.

Moujahid et al. [19] developed a CNN-based classification strategy by analysing pneumonia detected patients with X-ray lung images. Pre-processing, feature extraction, selection and classification were the steps for pneumonia diagnosis. In the first stage, the images had been prepared by reshaping or resizing the image to feed to the next stage. With the aid of the weight updation of the neurons the network gets trained. That can be validated by minimising the loss function and maximizing the accuracy rate. Here, the cross entropy was assimilated as a loss function, and softmax was inclined to function. The validated method attains accuracy, precision and f score of 96.81%, 91%, and 94% at 33 epochs.

Singh and Kolekar [23] developed deep learning strategy for COVID-19 diagnosis by utilising chest CT scan images in an edge cloud computing platform. The exemplary tune transfer learning model had been proposed and it had been processed under three stages. Pre-processing, fine-tuned Mobile Net V2 model, and classifier. In a Mobile Net V2 network model, convolutions were integrated in depth wise manner. The method had been differentiated into depth wise convolution and point wise convolution. The architecture had been incorporated with convolution layers, expansion layers and projection layers. The feature map dimension gets enhanced in the expansion layer and the high dimensional data had been converted into low dimensional data in the projection layer. The CT images were utilized for the experimental validation. The assessment metrics such as accuracy, precision, f score, sensitivity and specific attains a value of 96.4%, 94.6%, 96.5%, 94.4% and 93.2%, respectively. In addition to that the statistical significance test had also been examined.

ElpeltagyandSallam [7] had progressed automatic prediction of disease diagnosis from chest images by manipulating modified ResNet53. It has been implemented to differentiate the patients normal or abnormal. The enhancement in ResNet50 had been integrated with convolution layer, batch normalization layer and activation Relu layers. The layers as mentioned above had been inclined in the original architecture to extend the robustness of the developed model. Here, the weight updation in neural network had been handled by the adam optimizer with 1 e^{-4} learning rate and five epochs. The presented method attains the performance metrics of accuracy (97%), sensitivity (96.5%), specificity (95%) and precision (97%), respectively.

Achanta et al. [2] introduced hidden Markov model-based adaptive dynamic time warping to recognise the physically challenged persons. In that article, the author had developed a proficient gait recognition system to detect the abnormal condition of the person in a congested area and rendering desirable warn to the persons by keep track of them. The solution had easily gathered by hidden Markov model with the rid of data classification with several task nodes. The developed model attains the sensitivity, specificity and accuracy of 93.5%, 91% and 92.4% progressively. Achanta et al. [1] had evolved a gait detection approach in a wireless IoT system by using force sensitive resistor (FSR) sensor and wearable devices. Adaptive neuro fuzzy interference system (ANFIS) suggested that classification and ant colony optimization (ACO) were integrated to optimise efficient routing. The author originated both the

classification and optimization model to dispense the optimal routing path and to preserve the Qos requirements. The well organised developed model had been implemented in a matlab platform and it attains a recognition rate of 95% specifically.

2.1.1 Problem statement

From the surveys taken, it has been recognised that the conventional methods for pneumonia diagnosis classification are not so accurate and efficient. Nowadays, artificial intelligence methods have been broadly utilized to classify pneumonia X-ray images. But, the training process of these approaches needs more data to give acceptable performance. Thus, the complexity of these approaches has been increased. Also, their performance has been decreased due to the overfitting problem. The majority of the reviewed methods grab very long time to train and assemble the classification results. If the classification has to be performed with enormous images with large quantity of features, the classifier clutches high learning time to detect the normal and abnormal images. Several surveyed techniques are highly efficient and systematic in presuming better classification results, but those accurate techniques are computationally complex and impractical to implement in real-world. To compromise the immediate requirement of an accurate and efficient classification strategy and overcome the aforementioned issues, the efficient novel technique is proposed in this research. The selection of optimization strategy had emphasised the computational complexity in the neural network. Therefore, the appropriate selection of optimization with high convergence rate accomplishes effective classification results with low computational complexity. AMRFO algorithm is nominated and it diminishes the computational complexity under the essence of learning strategy with less error. The primary goal of the proposed work is to enhance the efficient classification outcome by mitigating the time and computational complexities. In this study a new hybrid network model is proposed to extract deep features by allowing self-attention mechanism in the network model also the loss function is optimized with the aid of meta-heuristic approaches.

3 Proposed methodology

In our work, the convolutional network is combined together with Mantaray optimization along with multiscale method to create a hybrid diagnosis for pneumonia. This methodology having three stages: pre-processing data, deep feature extraction and classification. The pre-processing of data is carried through the hybrid fuzzy color and stacking approach. This approach initially filter the data using fuzzy color method and the output image from fuzzy color is gone through the stacking approach to obtain restructured dataset with the best quality Hybrid multiscale convolution mantaray network is proposed to diminish the count of deep features without the loss of its original property. The multiscale feature extraction combined with self-attention module based convolutional neural network to minimize the network depth and extract more deep features with narrow configuration. The adaptive mantaray foraging optimization (AMRFO) algorithm is suggested to reduce the loss of function in the network by tuning the hyper-parameter. Finally, support vector regressor (SVR) based decision module identifies pneumonia by getting the selected feature sets as input. The framework of the developed method is illustrated in Fig. 1.
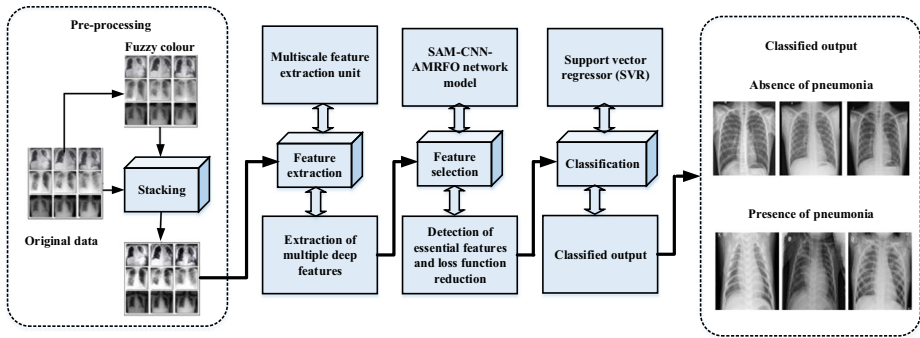
Fig. 1 Framework of proposed methodology

3.1 Computer-aided diagnosis

Computer-aided detection is a computer-based system that helps to analyse the disease through the medical image [18]. In deep learning, multilayer structure connects many middle layers from the hierarchical structure created from the first input layer to the last output layer [8]. In CAD diagnosis, data were collected as the image format and the data is collected without the loss of any information. The CAD diagnosis increases the accuracy of the identification of disease in certain area. In this proposed method the CAD helps to implement the diagnosis of pneumonia in human body.

3.2 Pre-processing data

3.2.1 Fuzzy color and stacking approach

The concept of fuzzy is established following its degree of accuracy. Fuzzy color approach plays a significant part in image processing analysis. Region of interest (ROI) sketches a certain area for analysis of image to get accurate information. The computer tomography (CT) images give multiple images with different angles to get to know about the problem. With the help of ROI, the process of analyzing the disease became easier. The ROI sketch the needed area of the CT images and proceed for the further process. The data are pre-processed using the hybrid fuzzy coloured and stacking approach. The original data is taken as input for the pre-processing data. In fuzzy color techniques, each information has three input variables RGB (red, green and blue). Fuzzy color technique converts the input data into blurry windows. The represented color image can be specified as follows:

$$I(u,v) = \{R(u,v), G(u,v), B(u,v)\} \tag{1}$$

where, $R(u,v)$, $G(u,v)$ and $B(u,v)$ specifies the red, green and blue channels. (u,v), denotes the pixel coordinate with $x \times y$ image size. The color attributes present in an image get augmented or improved due to the stretching of color channels. The process of stretching with respect to red color is given as follows:

$$R(u,v) \leftarrow \frac{R(u,v) - R_{\min}}{R_{\max} - R_{\min}} \tag{2}$$

where, $R_{\max} = \max R(u,v)$ and $R_{\min} = \min R(u,v)$ for all (u,v) respectively. Therefore the green and blue channels are stretched based on the same way. Inspite of this process, the image

get enhanced and the noises get removed. The intensity channel of an image can formulate the fuzzy dissimilarity histogram and it is defined as:

$$I = \{I(u, v) \mid 0 \leq u \leq X-1, 0 \leq v \leq Y-1\} \tag{3}$$

$I(u, v) = \rho_i$ signifies the intensity of the pixels in an image located at (u, v). The integer I ranges from 0 to 255. Membership degree can be computed based on the distance among the pixels and the window. The membership function can be stated as:

$$\chi(p, q) = \max\{1 - I(u, v) - I(p, q)/SD, 0\} \tag{4}$$

SD signifies the standard deviation. $p = u + i$, $q = v + j$; $i, j \in (-1, 0, 1)$. The fuzzy membership function can be sated by the involvement of membership values of neighbouring pixels. Therefore the fuzzy membership function based on fuzzy similarity index is stated as

$$\phi_{mf}(u, v) = \sum_{i=-1}^{1} \sum_{j=-1}^{1} \mu(u + i, v + j) \tag{5}$$

In the pre-processing, the elimination of noise takes place based on fuzzy color approach. The main aim of fuzzy colorapproachis the isolation of input data in to blurred windows. In this technique, the weight value of each pixels are ensemble as degree of membership. The noisy images from the dataset can be pre-processed with the fuzzy color technique.

3.2.2 Stacking techniques

The technique image stacking, deals with the image processing concept. Image stacking is a process of reconstructing the image to progress the quality of an image. The method signifies eradicating the noise from the input image by combining a minimum of two images and dividing it into two parts: background and overlay. Initially, the images are processed in background, and another one is acting as overlaying background image. Contrast, brightness, opacity and combining ratio are the important parameters of the two images. The stacking technique reconstructs the original data or image with the aid of fuzzy technique. The original and structured dataset are background and overlay. In the pre-processing stage, the noise are eradicated with fuzzy color technique and the original images or dataset get stacked to enhance the quality of a dataset. Then the stacked dataset is further utilized for feature extraction process.

3.3 Deep feature extraction

The hybrid multistate feature extraction approach with self-attention module (SAM) basedCNN is proposed to extract deep features and reduce the count of deep features without its original scheme. Feature extraction involves reducing the number of resources required to describe the dataset with a large set of data. Deep feature extraction may be done by extracting the high features from the image for precise identification of the image. The crucial set of features are selected according to neural networks. Multiscale feature extraction will extract features from different scales with various resolutions of frequency domain. After the extraction of features, self-attention based CNN is suggested to analyse the deep features in the training stage. CNN usually detect important features automatically without any human supervision. Adaptive mantaray foraging optimization (AMRF) is proposed to reduce the loss

function of the network. Hybridmultiscale SAM-CNN-AMRF network model is proposed to extract deep features, minimize the depth of the network with a selected set of features, and reduce the neural network's loss.

3.3.1 Multiscale feature extraction

The primary aim of this approach is to extract the features without any redundancy. Initially, the original image is divided into labelled datasets. Let us assume that the image is comprised of S sampling points and D data points. Partitioning of samples can be described as:

$$U(t) = \{(U_1, V_1), (U_2, V_2)....(U_N, V_N)\}, N = \frac{S}{D} \tag{6}$$

N specifies the number of samples. After the partitioning of samples, to compute the samples and compose the feature extraction approach, one dimensional kernel with various scales is established. The computation outcome can be transformed into two dimensional by concatenating. Therefore the expression can be stated as follows:

$$F = Concatenate \left[ReLU \left(\omega_1^1 U + b_1^1 \right)....ReLU \left(\omega_1^n U + b_1^n \right) \right] \tag{7}$$

$$O = down(F, m) \tag{8}$$

U signifies the sample elected from dataset, ω denotes the weight matrix with dissimilar scales, b denotes the bias vector, Re LU specifies the Rectified Linear unit is also stated as an activation function, $Concatenate(.)$ signifies the combination of one dimensional feature vectors and two dimensional feature maps by that the extracted features are analysed in a positive way. F signifies the two dimensional feature map. m and$down(.)$ signifies the scale and down sampling function of max pooling. Output feature map of the pooling layer can be denoted as$O$. The multi extracted features can be defined as follows:

$$Mean\ value\ \rho_1 = \frac{1}{i} \sum_{j=1}^{i} x(j) \tag{9}$$

$$Standard\ Deviation\ \rho_2 = \sqrt{\frac{1}{i-1} \sum_{j=1}^{i} \left[x(j) - \rho_1 \right]^2} \tag{10}$$

$$AbsoluteMean\ value\ \rho_3 = \frac{1}{i} \sum_{j=1}^{i} |x(j)| \tag{11}$$

$$waveform\ index\ \rho_4 = \frac{\rho_2}{\rho_3} \tag{12}$$

$$Maximum\,value\;\rho_5 = \max|x(j)| \tag{13}$$

$$Minimum\,value\;\rho_6 = \min|x(j)| \tag{14}$$

$$Peak{-}to{-}peak\,value\;\rho_7 = \rho_5{-}\rho_6 \tag{15}$$

$$Pulse\;index\;\rho_8 = \frac{\rho_5}{\rho_3} \tag{16}$$

$$Peak\;index\;\rho_9 = \frac{\rho_5}{\rho_2} \tag{17}$$

$$Square\,root\,amplitude\;\rho_{10} = \left(\frac{1}{i}\sum_{j=1}^{i}\sqrt{|x(j)|}\right)^2 \tag{18}$$

$$Margin\;index\;\rho_{11} = \frac{\rho_5}{\rho_{10}} \tag{19}$$

$$Skewness\;\rho_{12} = \frac{1}{i}\sum_{j=1}^{i}(x(j))^3 \tag{20}$$

$$Variance\;\rho_{13} = \frac{1}{i}\sum_{j=1}^{i}(x(j))^2 \tag{21}$$

$$Kurtosis\;\rho_{14} = \frac{1}{i}\sum_{j=1}^{i}(x(j))^4 \tag{22}$$

Multiscale feature extraction approach extracts multiple deep features and the extraction approach strengthen the capability of feature extraction. The extracted features are coordinated in a vector set format and it is expressed as follows:

$$FS = \left[\rho_1,\rho_2,\rho_3,\rho_4,\rho_5,\rho_6,\rho_7,\rho_8,\rho_9,\rho_{10},\rho_{11},\rho_{12},\rho_{13},\rho_{14}\right] \tag{23}$$

3.3.2 SAM-CNN-AMRF network model

The CNN has high accuracy and due to its high accuracy it is used for image classification and recognition. CNN models works in deep learning approach by train and test every input image

that passed through convolutional layer. Pooling layers are used as the filters in deep learning for the clear data and softmax function to classify the object. The pooling layer reduces the size of the spatial size to reduce the amount of parameter. The primary aim of the proposed network model is to diminish the number of essential deep features and keep down the network size with narrow configuration. Multiscale feature extraction approach with narrow receptive field can affect the convolutional layer when processing with high dimensional input data. Hence, the proposed model adds a self-attention module after every convolutional layer for capturing the long-range dependences by considering the weighted sum of features. Also, loss function of the proposed network is reduced using adaptive mantaray foraging optimization (AMRFO) algorithm by tuning the hyper parameters.

The proposed network model is emphasized with convolutional layer, self-attention layer and pooling layer. Convolutional layer is incorporated with convolutional kernels and it makes an output by manipulating the rectified linear unit (Re LU). Computational time get diminished with the rid of pooling layers. The information is transformed through three convolutional, self-attention, fully connected and max pooling layers in our approach. Both Re LU and softmax activation functions were utilized in convolutional and output layer. The expression of convolutional layers can be represented as:

$$U^i\left(\chi\right) = \max\left(0, b^i\left(\chi\right) + \sum K^{ij}(\chi)*v^i\left(\chi\right)\right) \tag{24}$$

where, $U^i\left(\chi\right)$, $v^i\left(\chi\right)$ and $b^i\left(\chi\right)$ specifies the input, output and bias activation map. $K^{ij}(\chi)$ specifies the kernel among input and output feature map. Self-attention mechanism is introduced to improve the concert of the CNN model. Therefore the attention value can be synthesized as:

$$V_p = \frac{1}{C\left(U_p\right)} \sum_q I\left(U_p, U_q\right) h\left(U_q\right) \tag{25}$$

U specifies the input fed to the self-attention block, V resembles the duet contribution of the input signal and also the output of the softmax function. P terms the recorded position and q resembles all position with different weights. h signifies the input feature set. I describes the intensity. Therefore the output of the attention layer is coordinated with feature and self-attention maps from prior convolutional layer. The parameter (α) in self-attention layer balance both components and it get learned through the training process.

The output size of the convolutional layers turns down by instigating the pooling layers. When the input is transformed through pooling layers, the deep set of features will get reduced and the maximum value is accumulated. Therefore the maximum pooling layer is defined as follows:

$$U^i_{jk} = \max\left(v^i_{js+m, Ks+n}\right) \tag{26}$$

U^i_{jk} signifies the neuron in the output activation map. The neurons in l layers are coupled with neurons in the output layer L-1. The coupling process is executed in the fully connected layers. In order to compute the weighted sum the connection among the L and L-1 layers is essential. The output of the fully connected layers is expressed as follows:

$$U^i_{jk} = \phi\left(u\right)\left(\sum_{i1}^{n^{(l-1)}} u^{(L-1)}\left(i\right).\omega^{(l)}\left(i, j\right) + b^{(l)}\left(j\right)\right) \tag{27}$$

where, $N^{(l-1)}$signifies the presence of neurons in the previous layers and $\phi(u)$ describes the activation function. Finally the softmax layer construct K dimensional reduced feature set. In neural network, the cross entropy loss is termed as a loss or error function. The loss function can be represented as follows:

$$L_f = -\sum_{i=1}^{n}\sum_{j=1}^{m} u_{ij} \log \left(\frac{e^{\omega_j^T u_i}}{\sum\limits_{m=1}^{K} e^{\omega_m^T u_i}} \right) \tag{28}$$

The SAM-CNN gets trained with vast and diverse range of input features. But, the network model comprises weight values that are changed randomly for every hidden layer, which highly influences the classification accuracy. To avoid this trouble, the optimization method is introduced to update the weights in each stage of the SAM-CNN network.

3.3.3 Adaptive Mantaray foraging optimization

The mantaray foraging optimization is developed by foraging of mantaray. The foraging techniques classified into three types based on the concentration of planktons. The mantaray also have a pair of vertical lobes which is in front of their mouth. Mantaray eats sea planktons and the foraging techniques are based on the plankton's concentration. Afitnessfunction is formulated to determine the optimality of the chosen weight factor. This can be given as follows:

$$Fit = \min(L_f) \tag{29}$$

The mantaray foraging optimization is processed with three process they are chain foraging, cyclone foraging and somersault foraging. The three unique and intelligent foraging strategies of manta rays can be identified as follows:

Chain foraging The chain foraging is defined as the mantaray lined up for foraging and moving on in the same line one after another. If the planktons are missed by one mantaray the next one will scooped them. By doing this cooperative way, they can able to collect most of the planktons. The position of every individual feature is identified and updated based on the following chain position update equation:

$$y_i^d(t+1) = \begin{cases} y_i^d(t) + (r+\alpha)\left(y_{best}^d(t) - y_i^d(t)\right) i = 1 \\ y_i^d(t) + r\left(y_{i-1}^d(t) - y_i^d(t)\right) + \alpha\left(y_{best}^d(t) - y_i^d(t)\right) i = 2, 3, \ldots N \end{cases} \tag{30}$$

$$\alpha = 2r\sqrt{|\log r|} \tag{31}$$

Here, $y_i^d$is the position of ith individual in the dth dimension, r is a random vector with ranges in between [0,1], $\alpha$is weight coefficient and y_{best}^d is high concentration of planktons.

Cyclone foraging When the concentration of planktons is very high, the huge amount of mantaray gathered together in the form of tail linkup with head of next one and move towards the food through spiral shape like path. The position of every foraging feature in the group is

updated based on the following position update equation:

$$-y_i^d(t+1) = \begin{cases} y_{best}^d + (r+\beta)\left(y_{best}^d(t) - y_i^d(t)\right) i = 1 \\ y_{best}^d + r\left(y_{i-1}^d(t) - y_i^d(t)\right) + \beta\left(x_{best}^d(t) - y_i^d(t)\right) i = 2,3,....N \end{cases} \qquad (32)$$

$$\beta = 2e^{r1\frac{T-t+1}{T}}\sin 2\pi r_1 \qquad (33)$$

Here, T represents the maximum iterations number, β is the weight coefficient, and r_1 denotes the random number in the range of [0, 1].

Each of individual get find a new position apart from the current position and that makes to obtain a global search. The mathematical expression is given as,

$$y_i^d(t+1) = \begin{cases} y_{rand}^d(t) + (r+\beta)\left(y_{rand}^d(t) - y_i^d(t)\right) i = 1 \\ y_{rand}^d(t) + r\left(y_{i-1}^d(t) - y_i^d(t)\right) + \beta\left(y_{rand}^d(t) - y_i^d(t)\right) i = 2,3,....N \end{cases} \qquad (34)$$

Here, y_{rand}^d denotes the random position.

Somersault foraging Somersault Foraging is defined as random, frequent, local, cylindrical movement of mantaray to foraging by individually. They do series of backward somersault which cycling around the planktons that go towards the mantaray. The mathematical expression is given as,

$$y_i^d(t+1) = y_i^d(t) + S\left(r_2 \times y_{best}^d - r_3 \times y_i^d(t)\right) \qquad (35)$$

Adaptive Mantaray Foraging Optimization is termed as an essential strategy to optimize the problems. In order to improve the prospects of an MFO, an adaptive mechanism is introduced. Initially the algorithm is improved by instigating opposition based learning technique. The utilization of learning technique will yields an effective outcome and it is given as follows:

$$\overline{y}_i^C = y_i^{max} + y_i^{min} - y_i^c \qquad (36)$$

where, $\overline{y}_i^C$ specifies the contrary location for y_i^c, y_i^{min} and y_i^{max} specifies the minimum and maximum constraints. In addition to learning technique, self-adaptive method is commenced to improve the MFO approach. This is mainly due to the variable size adaptation of an individual and it get changed by Eq. (37)

$$y_i^c = 10 \times P \qquad (37)$$

where, P denotes the parameter. Along with that, the updated individual size can be represented mathematically in Eq. (38).

$$\overline{y}_i^{C+} = round\left(x_i^c \times (1+\chi)\right) \qquad (38)$$

where, χ defines the random uniform distribution. The individual size gets increased or decreased if the χ is positive or negative. Moreover, the hybrid multiscale SAM-CNN with the help of AMRFO produces more accurate output with reduced error and it selects the optimal hyper-parameters.

3.4 Classification

The selected features with reduced dimensionality are directly given as an input to the support vector regressor (SVR) to identify the presence of pneumonia. SVR is termed as an supervised machine learning approach and it can be utilized for both regression and classification problems. SVR seeks the hyper plane to discrete the positive data from the negative data with high margin. The connections among the input and output variables can be recognized with the aid of structural risk minimization (SRM) norm and it can be computed based on the following equation:

$$Y = K(p) = \omega \phi(p) + m \tag{39}$$

$p = (p_1, p_2, \ldots .p_l)$ specifies the input data, $Y_b \in R^l$ resembles the resultant value, $\omega \in R^l$ denotes the weight factor, $m \in R$ represents the mathematical constant number and l denotes the data size. $\phi(p)$ specifies the irregular function corresponds to map the input features in an high dimensional feature space. To define the weight factor (ω) and mathematical constant number (m), the below formulae is emphasized in accordance with the concept of SRM:

$$Minimize \left[\frac{1}{2} \|\omega\|^2 + \rho \sum_{d=1}^{l} \left(\xi_d + \xi_d^* \right) \right]$$

$$subject to \begin{cases} y_d - (\omega\phi(p_d) + m_d) \leq \gamma + \xi_d \\ (\omega\phi(p_d) + m_d) - y_d \leq \gamma + \xi_d^* \\ \xi_d, \xi_d^* \geq 0 \end{cases} \tag{40}$$

where, ρ denotes the penalty function to maintain empirical risk and model flatness. ξ_d, ξ_d^* denotes the loose variables and γ represents the constant. The above equation describes the dual convex optimization problem and it can be solved with the aid of lagrangian function and it is defined as follows:

$$L\left(\omega, m, \xi_d, \xi_d^*, \beta_d, \beta_d^*\right)$$
$$= \frac{1}{2}\|\omega\|^2 + D\sum_{d=1}^{l}\xi_d + \xi_d^* - \sum_{d=1}^{l}\beta_d\left(\xi_d + \gamma - y_d + \omega\phi(p_d) + m\right) -$$
$$\sum_{d=1}^{l}\beta_d^*\left(\xi_d^* + \gamma - y_d - \omega\phi(p_d) - m\right) - \sum_{d=1}^{l}\left(\delta_d\xi_d + \delta_d^*\xi_d^*\right) \tag{41}$$

where, δ_d, δ_d^*, ξ_d and ξ_d^* are the lagrangian multipliers. Finally the regression function to classify the presence of pneumonia with the sake of SVR can be computed as follows:

$$K(p) = \sum_{d1}^{l} \left(\beta_d - \beta_d^*\right) k(p, p_d) + m \tag{42}$$

$k(p, p_d)$ denotes the kernel function. Finally, the SVR classifier classifies the pneumonia affected and non-affected classes.

4 Results and discussion

The most important aim of our proposed model is to progress the accuracy of pneumonia disease with the sake of chest X-ray images. In this section, we have verified the implementation of the proposed tactic with the existing detection methodologies. The efficacy of the

proposed method had been validated through kaggle dataset chest x-ray image (Pneumonia) dataset. The kaggle dataset had been used globally in many fields and it attains a better enriched outcome. The proposed pneumonia detection will be simulated in Python working platform. The dataset is comprised with 5863 x-ray images in JPEG format and it had chosen from the patients. Out of this, for a training phase 1341 normal and 3875 abnormal chest X-ray images are taken. And for testing phase, 234 normal and 390 abnormal X-ray images are taken to process. The dataset is developed with two types of classes (NORMAL and PNEUMONIA). The real dataset is comprised with three folders namely, training, testing and validation. Figure 2 shows the X-ray images with the presence and absence of Pneumonia.

The separation of input data in to blurred window can be performed by fuzzy color approach technique. In accordance with fuzzy colored approach, a digital image processing approach namely stacking technique is introduced to enhance the image quality in the dataset. This technique coordinated more than one images and reconstructed with various focal distances. Also it aims to eradicate the noises from the images. The hybrid fuzzy colored and the stacked image is shown in Fig. 3.

4.1 Training accuracy and training loss

The performance of proposed model is validated in terms of training accuracy and loss curve. In the training process both the parameters and hyper parameters are tuned. The Fig. 4 shows the testing and training accuracy under 25 epochs. In a training accuracy curve, our proposed model attains high accuracy of (97%) and signifies that the classification accuracy improves during training. Also, the training loss of our model attains very low when compared with
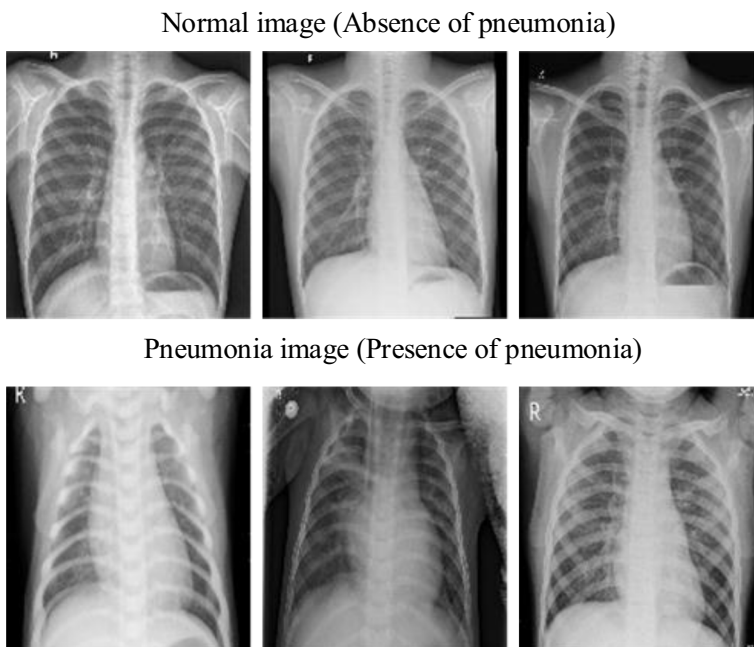
Normal image (Absence of pneumonia)

Pneumonia image (Presence of pneumonia)

Fig. 2 Presence and Absence of pneumonia (Sample images)
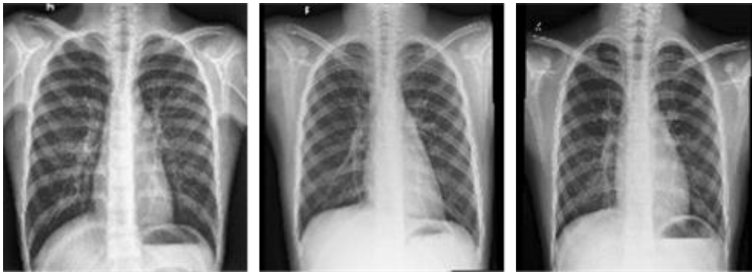
Stacked image

Fig. 3 Hybrid fuzzy colored and stacked image

existing classification strategies. From the figure, we can sense that our proposed strategy maximum training accuracy with minimum training loss.

4.2 Performance measure of quantitative analysis

This section deals with the suggested approach's quantitative performance metric under seven existing methodologies, namely VGG16, Dense Net, Resnet, DRE-Net, KNN, NN and SVM deep learning approaches. The metrics such as Accuracy, Precision, Recall (Sensitivity) and F-measures are evaluated. If the samples related to specific class are recognized appropriately for the classification approach, then the samples are placed in True positive (TP) indices. But if the samples related to some other classes are indentified exactly by the classifier then it is placed in True negative (TN) indices. Similarly, False positive (FP) and False negative (FN) specifies the samples are not correctly predicted by the classifier. The performance metrics are defined as follows:

4.2.1 Accuracy

The ability of a predictor is measured to identify the samples correctly, whether it is positive or not. Accuracy is an essential metric to validate the proposed detection methodology's efficacy compared with existing methodologies. The effectiveness of the proposed method had been proven through existing methods such as VGG16, Dense Net, Resnet, DRE-Net, KNN, NN and SVM. The mathematical expression of accuracy is termed as follows:
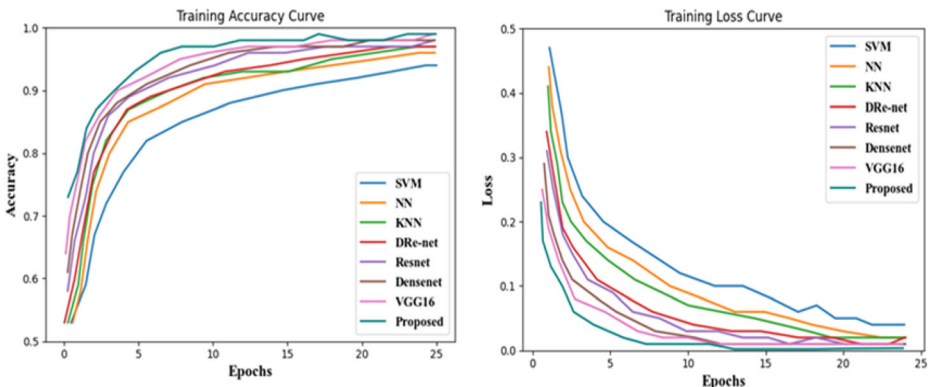

Fig. 4 Training accuracy and loss curve

$$A_y = \frac{Tp + Tn}{Tp + Tn + Fp + Fn} \qquad (43)$$

The accuracy measure of detection methodology in case of presence and absence of pneumonia plays an imperative role in the performance analysis metrics. TP and TN define the number of truly identified pneumonia and normal cases. The chest X-ray images are subjected to recognize the healthy and wealthy images with high promising accuracy. Figure 5 illustrates the graphical representation of accuracy. The proposed method achieves high outcome when analyzed with existing techniques. The proposed model attains (97%) and the existing VGG16, Dense Net, ResNet, DRE-Net, KNN, NN and SVM attains 91%, 87%, 90%, 95%, 83%,84% and 84% respectively.

4.2.2 Precision

Precision is also entitled as past predictive value (PPV) and it defined as the ratio of true positive rate to all positive values. And also it evaluates how accurate this model is. Precision can be defined as the classified images under labels. It can be summarized as the sum of labelled images to the number of classified images. The precision can be calculated by Eq. (44).

$$P_n = \frac{Tp}{Tp + Fp} \qquad (44)$$

The true positive values and false positive values authenticate the overall accuracy of the system. Precision regards that the accurately classified positives beyond all positive precisions. The false positive defines the number of normal cases can be misclassified as pneumonia cases. The graphical representation of precision metric is depicted in Fig. 6 with succeeding approaches. From Fig. 6 we come to know our proposed model attains high value of (95%) and the prevailed techniques of VGG16, Dense Net, Resnet, DRE-Net, KNN, NN and SVM attains 80%, 76%, 81%, 79%, 83%, 74% and 92% respectively.
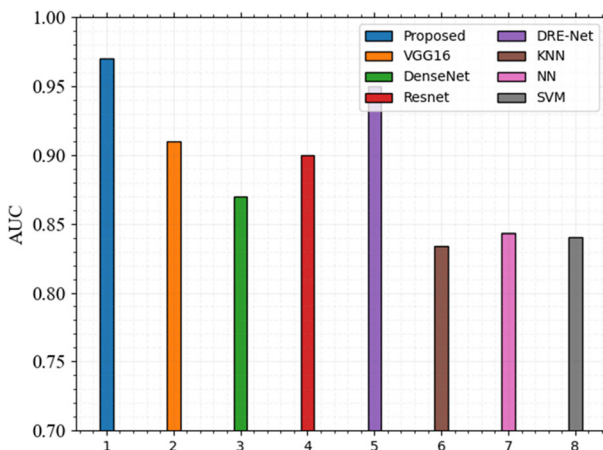

Fig. 5 Graphical view of Accuracy

4.2.3 Recall (sensitivity)

Recall measures the actual positives of the model and also it is emphasized to determine the accuracy of the proposed system model. The measure of sensitivity is also entitled as true positive rate (TPR). Sensitivity measures the capability of recognizing the positive samples and it is calculated by the below Eq. (45):

$$R_l = \frac{Tp}{Tp + Fn} \tag{45}$$

Where F_n signifies the false positives, which mean the number of pneumonia images incorrectly recognized as healthy images. The occurrence of error in a classification approach is rectified via this metric. The graphical representation of recall is demonstrated in below figure. From Fig. 7 it has been shown that our proposed tactic reveals a better recall value than the other techniques compared with it. The recall value of proposed method is 97%, and the existing method of VGG16, Dense Net, Resnet, DRE-Net, KNN, NN and SVM attains 89%, 93%, 93%, 96%, 84%, 75% and 83% respectively.

4.2.4 F-measure

F-measure computes the harmonic mean among the precision and sensitivity. The below figure illustrate the F-measure. The proposed method attains 96% and the existing methods such as VGG16, Dense Net, ResNet, DRE-Net, KNN, NN and SVM attains 84%, 83%, 86%, 87%, 84%, 75% and 90%respectively. The mathematical expression can be defined as follows:

$$F-measure = \frac{2 \times \text{Precision} \times Sensitivity\,(\text{Re}call)}{\text{Pr}ecision + Sensitivity} \tag{46}$$

In the above mentioned metrics and equation, when the classification takes place under both the healthy and pneumonia patients TP, TN, FP and FN are defined. The count of pneumonia images recognized as pneumonia signifies the true positive rate. If the normal or healthy images can be recognized as normal signifies the true negative rate. The normal image which
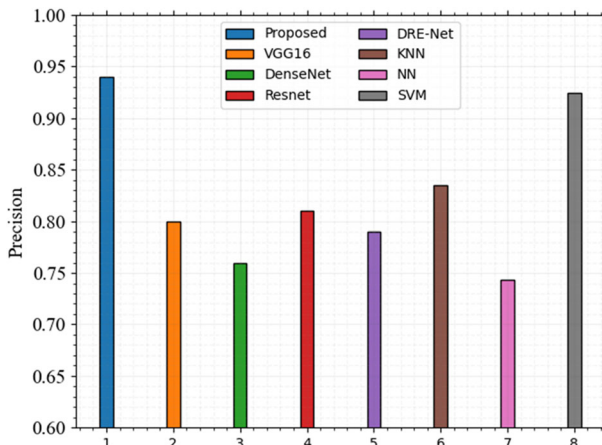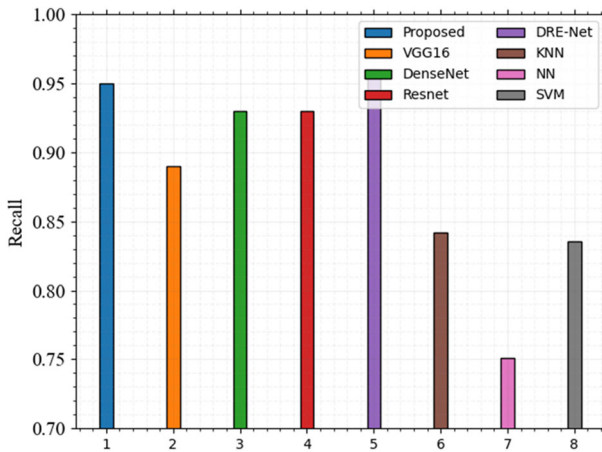

**Fig. 6** Graphical view of Precision

Fig. 7 Graphical view of Recall

had picked incorrectly as pneumonia image is termed as false positive rate. Finally, the false negative rate specifies the pneumonia images can be incorrectly detected as healthy. Based on our proposed approach, we have achieved pneumonia detection accuracy, specificity, sensitivity and F-measure of 97%, 95%, 97% and 96% respectively.

　The experiments carried out in this section is emphasised with three stages, namely preprocessing, feature extraction and classification. Deep learning has reformed automatic disease diagnosis and management by meticulously analyzing, recognising and classifying patterns in medical images. With the advent of artificial intelligence (AI), the computer-aided diagnosis (CAD) tools have also remarkable growth. In first stage, the fuzzy color technique and stacking techniques are emphasised for the elimination of unnecessary noises and to escalate the quality of dataset. The fuzzy concept is instigated by concerning its degree of accuracy. And, this technique is supposed to distinct the input data in to blurred windows. Stacking approach is initiated to upsurge the quality of the presence of image in the dataset. The hybrid fuzzy color and the reconstructed image are illustrated in Fig. 3. In the second stage, feature extraction is processed with the aid of neural networks. Hybrid multistate feature extraction unit is intended for the extraction of deep features. The features are extracted with various resolutions and in this stage, the depth of the network gets minimised, loss function gets reduced and deep features are extracted. In this stage, the most relevant features are selected. In third stage, the selected features are fed to the classifier to differentiate the presence and absence of pneumonia. The input and output image is illustrated in Fig. 2. The performance of the proposed network model is validated in terms of accuracy, precision, recall and F measure and it is shown through the graphical representations of Figs. 5, 6, 7 and 8, respectively.

　Patients with health issues such as pulmonary edema, bleeding, atelectasis, lung cancer or surgical interventions will make the pneumonia diagnosis more complex. Several methods had developed to classify the healthy and un-healthy X-ray images and in this research, the proposed method is evaluated with seven DL approaches, namely VGG16, Dense Net, ResNet, DRE-Net, KNN, NN and SVM, to validate the efficacy of the developed method. The existing methods are suffered due to some major reasons. The KNN method tries to identify the K-nearest neighbour, but the points nearest to are random and so, in this case, the method gets failed. The KNN classifier
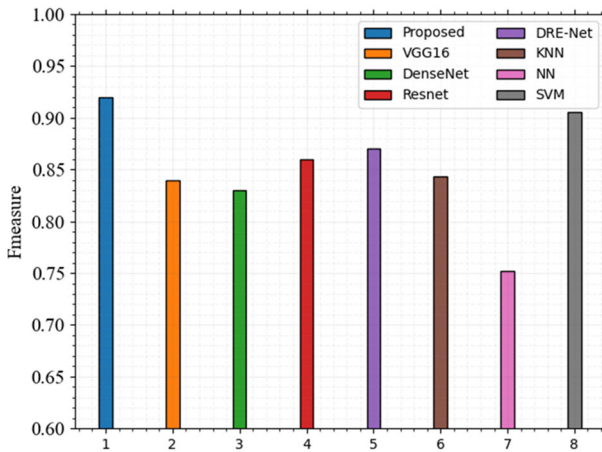
Fig. 8 Graphical view of Recall

had not potentially significant with large dataset and high dimensions. It needs standardization and normalization (feature scaling) before applying KNN to the data. If it fails to augment the feature scaling, the disease gets misclassified and it attains a very low accuracy of 83%, precision of 83%, recall of 84% and f score 84% respectively. The NN attains an accuracy, precision, recall and f score of 84%, 74%, 75% and 75% consecutively. The method SVM is also validated with the developed model. The kernel function in SVM makes strengthen to solve the complex issues. Still, the selection of best kernel function was difficult and also the fine tuning of hyper-parameters makes the classification accuracy demand. It procures 84%, precision of 92%, recall of 83% and f score of 90% progressively. The VGG 16 network is well suitable for learning and easy to implement. VGG-16 transfer learning attains a good result, particularly in a fine tuning task and also it yields high accuracy coverage. But, the network is insisted with enormous weights parameter with high interference time. So it makes the pneumonia diagnosis very complex. Therefore, this method attains an accuracy of 91%, precision of 80%, recall of 89% and f score of 84% accordingly. The existing DRE-Net had also been compared with the developed network model and it attains an accuracy of 95%, precision of 79%, recall of 96% and f-score 87% accordingly. In the existing ResNet, the preceding feature from the chest X-ray images was reused by ResNet with identity shortcuts. The features had safeguarded by the elimination of residual blocks. But the identity shortcut had degraded with collapsing domain issues that highly diminish the network's learning capacity. The classification accuracy gets varied and yields an accuracy of 90%, precision 81%, recall 93% and f score 86%, respectively. The subsequent layers in a Dense Net adopt dense concatenation for preserving the features, but it requires large memory. Compared with the proposed network model, the approach is Dense Net, and this network is emphasised with enormous connection layers. And these excessive connections lead to diminishing the computational efficiency and parameter efficiency and also, consumes enormous memory during the process of training. This method attains an accuracy of 87%, precision of 76%, recallof 93% and f score 83%respectively. Apart from existing methods, the proposed method yields an accuracy of (97%), precision of (95%), recall of (97%) and f score of (96%) correspondingly.

4.3 Confusion matrix analysis

The comparison of the actual value with the predicted value is demonstrated through confusion matrix analysis. The measure of classification analysis of normal and pneumonia cases can be classified with confusion matrices. Figure 9 depicts the confusion matrix of the proposed model. In figure, it is seen that our proposed classification methodology perfectly detects pneumonia and normal images. The testing phase is comprised with 234 normal images and 390 abnormal (pneumonia) images. For a healthy image, 229 images are predicted normally and only five images are misclassified and for abnormal images, 380 samples are correctly predicted as pneumonia and only ten images are misclassified as normal. From this confusion matrix analysis we come to know the overall accuracy is around 98% respectively. For a combined fuzzy and stacking approach in a healthy image, 227 images are predicted normal and rest of 7 images are misclassified as pneumonia. For pneumonia images, 377 images are correctly predicted as pneumonia and only 13 images are classified as healthy images.

4.4 Comparison of proposed framework with state-of–the-art techniques

The proposed scheme is seen to dominate the other approaches in terms of accuracy and other performance metrics from the simulations carried out. The effectiveness and efficiency of the proposed approach has been compared with the other state-of-the-art models relevant to pneumonia classification. It is clearly evident from the simulations that the proposed framework provided optimal results compared to other classification approaches. Rahman et al. [22] presented four different deep CNN models to detect bacterial and viral pneumonia under the influence of x-ray images. Here, the author reported the performance in three cases: normal, viral, and bacterial pneumonia. The presence of neurons in the filters is associated with local patches to enhance the spatial structure. Various pre-trained method had been emphasised to inflate the classification accuracy perhaps the radiographic findings destroy the benefit of the work. The scheme resulted in the accuracy of normal and pneumonia images, bacterial and viral pneumonia images and normal, bacterial, and viral pneumonia with better performance. The DNN for pneumonia localization [24] had been validated in the largest clinical database for pneumonia diagnosis. Here, the DNN was utilized to forecast pneumonia affected areas. Prediction of affected area will make support for a doctor to make a pneumonia diagnosis. Two CNN networks, namely RetinaNet and Mask R-CNN, are introduced for detection and
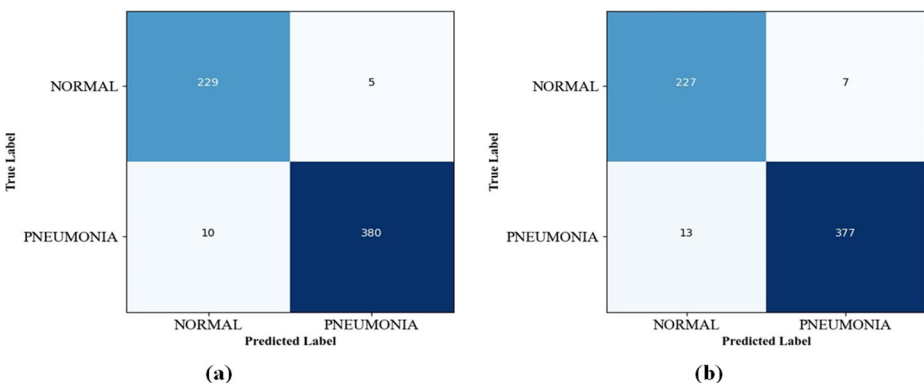
Fig. 9 Confusion matrix analysis (**a**) Pneumonia classification, (**b**) Combined fuzzy and stacking technique

localization. The method RetinaNet surpasses the accuracy and this method had the ability to predict the probability of an object. The each layer in RetinaNet generates the dense candidate frames and this leads to a negative number of samples. The Mask R-CNN is not much accurate and the quality improvement makes a challenging one in Mask R-CNN. Hence, the adjacent pixels in an image have identical image information, leading to upsurge or decline in recognition and classification accuracy. The method showed a precision, recall and f score of 75%, 79% and 77% respectively. The viral pneumonia screening [30] using the CAAD model was adapted and the model is assimilated with feature extractor, an anomaly detection module, and a confidence prediction module. The shared testing and training data are identical and often occur class-imbalance issues. Accordingly, the classifier capitulates indigent sensitivity performance and this leads to misdiagnosing healthy control. The experimental analysis of the CAAD model yields an outcome of AUC of 83.61%, accuracy 72.77%, sensitivity 71.70% and specificity 73.83% progressively.

Khan et al. [14] developed a deep neural network to detect and diagnose normal and abnormal images. The author had presented CoroNet for the detection of images. The presented model showed an accuracy of 89.6% in the case of COVID vs Pneumonia, bacterial vs pneumonia, viral vs Normal. The method had attained an overall accuracy of 90%. CoroNet can still be favourable for radiologists and health experts to capture extensive understandings into expository regard associated with COVID-19 cases. Ibrahim et al. [11] had presented pneumonia classification with the aid of four way classification approach. Several researchers had developed two ways and three way classification but here, the author had introduced four way classifications using AlexNet architecture and it showed an overall accuracy of 93%, sensitivity of 89% and sensitivity of 98%, respectively. Demir et al. [6] had defined CNN approach for classifying lung diseases. Here, the lung sound signals were transformed into spectrogram images based on time frequency methods (short time Fouriertransform (STFT)). CNN network and SVM classifier had intended for the extraction of features and classification. The method had turned out a result of 65% and 63% of accuracy progressively.

The pneumonia diagnosis concerning machine learning approaches had introduced by Yee and Raymond [28] to train and test the data. The author had defined KNN, NN and SVM strategy, but these methods diminished the stability and accuracy of classification. KNN is a lazy learner, which means it had no training period to train the data. In addition to this, it requires large memory to accumulate the training data and also it is computationally expensive so that the training accuracy of the presented work had declined and it attains an accuracy, precision, recall and f score of (83%), (83%),(84%) and (84%) consecutively. NN generates a potential outcome when compared to machine learning techniques and also it generates less error. And also, it trains the data in a parallel way; by this, the classification had performed earlier, but it is computationally expensive. The NN method yields an accuracy, precision, recall and f score of (84%), (74%), (75%) and (75%) consecutively. SVM classifier had been well appropriate with unstructed and semi-structured data like text or images. SVM is integrated with linear, polynomial, radial basis, and sigmoid kernel functions. SVM approach procures extensive training time while executing with large dataset. Therefore, the linear kernel yields a sensitivity of 83.5% and the rest of kernels had attained a sensitivity of 63%, 73% and 76% respectively. Apart from ML approaches, DL approached had developed to diagnosis the COVID-19 [26]. Here, the DRE-net architecture had presented by the author for the COVID-19 diagnosis. Here, VGG16, Dense Net and ResNet had validated with DRE-net architecture. The method of VGG16 yields an accuracy of (91%), precision of (80%), recall of (89%), and f score of (84%) consecutively. The method of Dense Net attains an accuracy, precision, recall

Table 1 Performance comparison of state-of-the-art techniques with proposed framework

Methodology	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Proposed	**97%**	**95%**	**97%**	**96%**
VGG16	91%	80%	89%	84%
Dense Net	87%	76%	93%	83%
ResNet	90%	81%	93%	86%
DRE-Net	95%	79%	96%	87%
KNN	83%	83%	84%	84%
NN	84%	74%	75%	75%
SVM	84%	92%	83%	90%
Dense Net 201	93.2%	93.7%	93.2%	93.5%
ResNet18	87.7%	87.5%	88%	90%
Alex Net	88.4%	88%	83%	88%
Mask R-CNN	–	75%	79%	77%
DRE-Net	95%	79%	96%	87%
CNN	96%	91%	–	94%
Mobile Net V2 model	96%	91%	97%	94%

and f score of 87%, 76%, 93% and 83% respectively. The performance measures of accuracy, precision, recall and f score attains 90%, 81%, 93% and 86% with the intention of ResNet. The author presented DRE-net method and it yields an accuracy of (95%), precision of (79%), recall of (96%) and f score of (87%) respectively.

CNN based transfer learning approach presented by [19] for pneumonia diagnosis in a lung images. The CNN consumes large amount of data to enhance the classification accuracy and it will not encode the position and the orientation of the image. The results had turned out with accuracy of (96%), precision of (91%) and f score of (94%) progressively. This scheme is mostly equivalent with the proposed tactic but it is computationally complex. The identical approach for COVID-19 diagnosis had presented by [23] using fine-tuned Mobile Net V2 model. The Mobile Net V2 had integrated with inverted residual block. Compared to other methods it holds short time period to train the data. The experimental results turned out that the presented method yields an accuracy of 96.81%, precision of 91%, recall of 97% and f score of 94% respectively. Another horrible method for COVID-19 diagnosis is ResNet 50 [7]. The method had presented by the author to detect healthy and un-healthy patients. But it requires prolonged time to train the data, but it is practically insoluble in real time applications. The developed approach showed an improved performance in terms of accuracy, sensitivity, specificity and precision independently. The comparison of the state-of-the-art-methods with the proposed method is tabulated in Table 1:

5 Conclusion

Pneumonia is one of the respiratory infections and a vulnerable disease that arises in the lungs. It affects humans in any time either by bacterial or viral infection. In this article, a hybrid multiscale convolutional mantaray network model is developed to upsurge the overall accuracy and diminish the classification error. The reason behind such success is that DL techniques do not depend on manual handcrafted features, but these algorithms learn features spontaneously from data itself. At present, the convolutional neural networks (CNNs) are

known to perform better in both general and medical image-based applications. These networks can be efficiently trained to detect normal and diseased patterns from the radiographic scans. In this research, hybrid multiscale convolutional mantaray network model is developed to upsurge the overall accuracy and diminish the classification error. Our approach is processed under three main sub stages, namely pre-processing, feature extraction and classification. Theraw data is filtered through the hybird fuzzy color and stacked approach to eradicate unwanted noises and enhance the image quality. In an extraction phase, from the stacked dataset the features are extracted with the rid of hybrid multiscale feature extraction unit and the size of the network and features are diminished with the sake of SAM based CNN. Finally, in the CNN model, the SVR is incorporated to classify the healthy and pneumonia images. The error is classification approach is tackled with AMRFO approach. The quantitative and confusion matrix results proclaim the preferable outcome for the proposed tactic compared with existing methods. The performance of the proposed model attains an accuracy of 97%, precision 95%, recall 97% and F-score 96%. In the future, this study had enhanced with various deep learning solutions to detect and identify the affected area in the radiographic scans. As well as in future, we plan to develop structuring approaches to enhance the quality of the dataset.

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

References

1. Achanta SDM, Karthikeyan T (2020) A wireless IOT system towards gait detection technique using FSR sensor and wearable IOT devices. Int J Intell Unmanned Syst 8(1):43–54. https://doi.org/10.1108/IJIUS-01-2019-0005
2. Achanta SDM, Karthikeyan T, Vinothkanna R (2019) A novel hidden Markov model-based adaptive dynamic time warping (HMDTW) gait analysis for identifying physically challenged persons. Soft Comput 23:8359–8366
3. Asuntha A, Srinivasan A (2020) Deep learning for lung cancer detection and classification. Multimed Tools Appl 79(11):7731–7762
4. Chen H-J, Ruan S-J, Huang S-W, Peng Y-T (2020) Lung X-ray segmentation using deep convolutional neural networks on contrast-enhanced Binarized images. Mathematics 8(4):545
5. Chouhan V, Singh SK, Khamparia A, Gupta D, Tiwari P, Moreira C, Damaševičius R, De Albuquerque VHC (2020) A novel transfer learning based approach for pneumonia detection in chest X-ray images. Appl Sci 10(2):559
6. Demir F, Sengur A, Bajaj V (2020) Convolutional neural networks based efficient approach for classification of lung diseases. Health Inform Sci Syst 8(1):1–8
7. Elpeltagy M, Sallam H (2021) Automatic prediction of COVID− 19 from chest images using modified ResNet50. Multimedia tools and applications 1-13.
8. Fujita H (2020) AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. RadiolPhysTechnol 13:6–19. https://doi.org/10.1007/s12194-019-00552-4
9. Hashmi MF, Katiyar S, Keskar AG, Bokde ND, Geem ZW (2020) Efficient pneumonia detection in chest xray images using deep transfer learning. Diagnostics 10(6):417
10. Hu Q, deSouza LF, Holanda GB, Alves SSA, dos Silva FH, Han T, Filho PPR (2020) An effective approach for CT lung segmentation using mask region-based convolutional neural networks. Artif Intell Med 103:101792
11. Ibrahim AU, Ozsoz M, Serte S, Al-Turjman F, Yakoi PS (2021) Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cogn Comput:1–13
12. Jain R, Nagrath P, Kataria G, Kaushik VS, Hemanth DJ (2020) Pneumonia detection in chest X-ray images using convolutional neural networks and transfer learning. Measurement 165:108046
13. Jaiswal AK, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJPC (2019) Identifying pneumonia in chest X-rays: a deep learning approach. Measurement 145:511–518

14. 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 Prog Biomed 196:105581

15. Li Y, Zhang Z, Dai C, Dong Q, Badrigilan S (2020) Accuracy of deep learning for automated detection of pneumonia using chest X-ray images: a systematic review and meta-analysis. Comput Biol Med 123: 103898

16. Liang G, Zheng L (2020) A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Comput Methods Prog Biomed 187:104964

17. Mahmud T, Rahman M, Fattah SA (2020) CovXNet: a multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. Comput Biol Med 122:103869

18. Moon WK, Lee Y-W, Ke H-H, Lee SH, Huang C-S, Chang R-F (2020) Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks. Comput Methods Prog Biomed 190:05361. https://doi.org/10.1016/j.cmpb.2020.105361

19. Moujahid H, Cherradi B, El Gannour O, Bahatti L, Terrada O, Hamida S (2020) Convolutional neural network based classification of patients with pneumonia using X-ray lung images. Transfer 2(99.41):16

20. Ni Q, Sun ZY, Qi L, Chen W, Yang Y, Wang L, Zhang X, Yang L, Fang Y, Xing Z, Zhou Z, Yu Y, Lu GM, Zhang LJ (2020) A deep learning approach to characterize 2019 coronavirus disease (COVID-19) pneumonia in chest CT images. Eur Radiol 30(12):6517–6527

21. Qin J, Pan W, Xiang X, Tan Y, Hou G (2020) A biological image classification method based on improved CNN. Ecol Inform 58:101093

22. Rahman T, Chowdhury MEH, Khandakar A, Islam KR, Islam KF, Mahbub ZB, Kadir MA, Kashem S (2020) Transfer learning with deep convolutional neural network (CNN) for pneumonia detection using chest X-ray. Appl Sci 10(9):3233

23. Singh VK, Kolekar MH (2021) Deep learning empowered COVID-19 diagnosis using chest CT scan images for collaborative edge-cloud computing platform. Multimedia Tools Appl:1–28

24. Sirazitdinov I, Kholiavchenko M, Mustafaev T, Yixuan Y, Kuleev R, Ibragimov B (2019) Deep neural network ensemble for pneumonia localization from a large-scale chest x-ray database. Comput Electrical Eng 78:388–399

25. Skourt BA, El Hassani A, Majda A (2018) Lung CT image segmentation using deep neural networks. Procedia Comput Sci 127:109–113

26. Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, Chen J, Wang R, Zhao H, Chong Y, Shen J, Zha Y, Yang Y (2021) Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. IEEE/ACM Trans Comput Biol Bioinform 18:2775–2780

27. Varela-Santos S, Melin P (2020) Classification of X-ray images for pneumonia detection using texture features and neural networks. In: Intuitionistic and Type-2 fuzzy logic enhancements in neural and optimization algorithms: theory and applications. Springer, Cham, pp 237–253

28. Yee SLK, Raymond WJK (2020) Pneumonia diagnosis using chest X-ray images and machine learning. In proceedings of the 2020 10th international conference on biomedical engineering and technology, 101-105.

29. Yue Z, Ma L, Zhang R (2020) Comparison and validation of deep learning models for the diagnosis of pneumonia. Comput Intell Neurosci 2020:1–8

30. Zhang J, Xie Y, Pang G, Liao Z, Verjans J, Li W, Sun Z, He J, Li Y, Shen C, Xia Y (2020) Viral pneumonia screening on chest X-rays using confidence-aware anomaly detection. IEEE Trans Med Imaging 40:879–890