

RESEARCH

Open Access



# Adaptive Mish activation and ranger optimizer-based SEA-ResNet50 model with explainable AI for multiclass classification of COVID-19 chest X-ray images

S. R. Sannasi Chakravarthy<sup>1</sup>, N. Bharanidharan<sup>2</sup>, C. Vinothini<sup>3</sup>, Venkatesan Vinoth Kumar<sup>2</sup>, T. R. Mahesh<sup>4</sup> and Suresh Guluwadi<sup>5\*</sup>

## Abstract

A recent global health crisis, COVID-19 is a significant global health crisis that has profoundly affected lifestyles. The detection of such diseases from similar thoracic anomalies using medical images is a challenging task. Thus, the requirement of an end-to-end automated system is vastly necessary in clinical treatments. In this way, the work proposes a Squeeze-and-Excitation Attention-based ResNet50 (SEA-ResNet50) model for detecting COVID-19 utilizing chest X-ray data. Here, the idea lies in improving the residual units of ResNet50 using the squeeze-and-excitation attention mechanism. For further enhancement, the Ranger optimizer and adaptive Mish activation function are employed to improve the feature learning of the SEA-ResNet50 model. For evaluation, two publicly available COVID-19 radiographic datasets are utilized. The chest X-ray input images are augmented during experimentation for robust evaluation against four output classes namely normal, pneumonia, lung opacity, and COVID-19. Then a comparative study is done for the SEA-ResNet50 model against VGG-16, Xception, ResNet18, ResNet50, and DenseNet121 architectures. The proposed framework of SEA-ResNet50 together with the Ranger optimizer and adaptive Mish activation provided maximum classification accuracies of 98.38% (multiclass) and 99.29% (binary classification) as compared with the existing CNN architectures. The proposed method achieved the highest Kappa validation scores of 0.975 (multiclass) and 0.98 (binary classification) over others. Furthermore, the visualization of the saliency maps of the abnormal regions is represented using the explainable artificial intelligence (XAI) model, thereby enhancing interpretability in disease diagnosis.

**Keywords** X-ray, COVID-19, Transfer learning, Explainable artificial intelligence, Deep-learning, Attention mechanism

\*Correspondence:

Suresh Guluwadi  
[suresh.guluwadi@astu.edu.et](mailto:suresh.guluwadi@astu.edu.et)

Full list of author information is available at the end of the article



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

## Introduction

Being a contagious and deadly disease, COVID-19 has a hazardous impact on the human respiratory system. Due to this, on March 11, 2020, the World Health Organization (WHO) declared this illness as a global pandemic severity with the origin of Wuhan, China [1]. Then the disease was recognized as a corona virus with around 75% similar to the SARS variant [1]. Additionally, the disease was almost similar to the bat corona virus that was recognized on earlier time of 2020. The symptoms of the disease vary from lower to higher risk of severity and cause multi-organ impairment. The resultant damages are associated with respiratory disorders such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) [2]. The initial indications of this disease start with a heavy cough, fever, respiratory issues, and breathing problems [2]. Until now, clinicians have used the following methods for diagnosing COVID-19 victims: the first one involves clinical tests namely reverse transcription polymerase chain reaction (RT-PCR) and the second one involves the utilization of antigen testing [3]. Due to more false reports provided by the antigen tests, RT-PCR has become the wider and more popular one for diagnosis. However, its requirement involves experienced clinicians with extensive laboratory work for obtaining and analyzing the results [4]. This makes this RT-PCR as a time-consuming and expensive test. Thus making a demand among researchers for addressing this concern [4].

Medical imaging and analysis using Artificial Intelligence (AI) algorithms have become an efficient tool for several biomedical problems. As an important point, the integration of AI methods into existing clinical workflows for COVID-19 classification enhances diagnostic capabilities, optimizes resource utilization, and improves patient care delivery. AI plays a crucial role in combating the COVID-19 pandemic and alleviating its impact on healthcare systems by complementing human expertise with advanced computational techniques. In this way, several researchers have been enthused with the development of AI-assisted Computer-Aided Diagnosis (CAD) frameworks for the robust classification of COVID-19. On the other hand, CAD frameworks with fewer false reports are needed for the diagnosis utilizing medical image databases. For this, radiographic images such as Computed Tomography (CT) and chest X-rays have proven to be more promising modalities. Extensive research and investigation have been conducted to portray the significance of CT image analysis for COVID-19 classification [5]. But, because of the good availability, lower cost, and the involvement of less radiation, chest X-ray images are preferred for detecting COVID-19 [5]. This makes this

study to utilize public datasets that contain chest X-ray images with multiple classes for the robust COVID-19 classification.

The work proposes an automated residual deep-learning architecture based on the Squeeze-and-Excitation Attention mechanism. And the model is named as SEA-ResNet50 utilizing the Ranger optimizer and adaptive Mish activation function for multiclass classification of COVID-19 severities. For comparative experimentation, the work explored distinct transfer learning models namely VGG-16, Xception, ResNet18, ResNet50, and DenseNet121 architectures. Moreover, the input images are preprocessed and augmented for better evaluation of the proposed model. Figure 1 illustrates the aforementioned points regarding the proposed work for the multiclass classification of chest X-ray images.

As shown in Fig. 1, this paper makes the following contributions:

- Proposes an end-to-end CAD framework based on the improved ResNet50 (SEA-ResNet50) model.
- Implements the backbone of the architecture using the Squeeze-and-Excitation Attention (SEA) mechanism, with the inclusion of the Ranger optimizer and adaptive Mish activation function.
- Evaluates the study using two COVID-19 chest X-ray image datasets.
- Conducts comparative experimentation with seven popular transfer learning architectures for multiclass classification of COVID-19 severity.
- Represents the visualization of the saliency maps of the abnormal regions in the chest X-ray images using the explainable artificial intelligence (XAI) model.

The rest of the sections of the paper are structured as follows: "[Related works](#)" section will signify the background study and its related works, "[Materials and methods](#)" section will illustrate the materials and methods used for the proposed study such as dataset selection, composition, and preprocessing techniques, "[Transfer learning models for multiclass classification of COVID-19 chest X-ray images](#)" section will discuss about the details of the transfer learning models, "[Proposed Squeeze-and-Excitation Attention—based ResNet50 \(SEA-ResNet50\) model for COVID-19 detection](#)" section will provide a detailed discussion on the proposed SEA-ResNet50 for COVID-19 classification, "[Experimental outcomes and analysis](#)" section will illustrate the experimental investigation followed by comparative analysis, and finally the conclusion and future extension of the proposed methodology will be detailed in "[Conclusion and future work](#)" section.

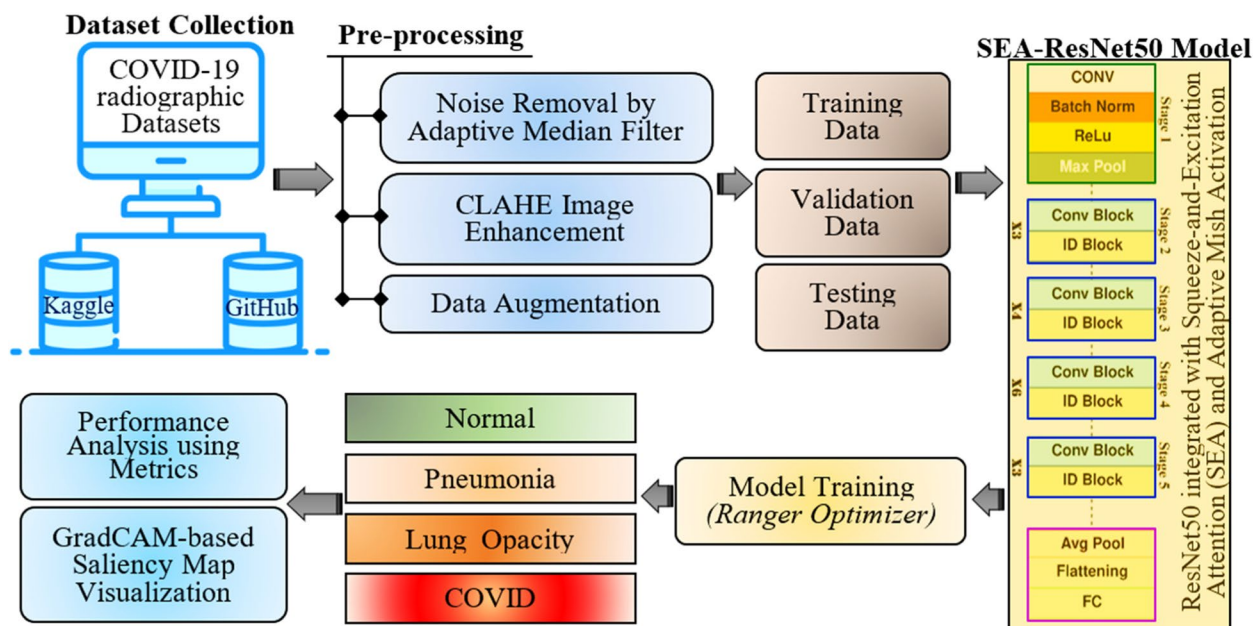


Fig. 1 Workflow for COVID-19 (Chest X-ray images) classification

**Related works**

As reviewed in the previous section, COVID-19 is a cruel and deadly disease affecting several human lives all around the world. Several researchers have extensively proposed different approaches for the multiclass classification of COVID-19. Out of which, some of the significant contributions are discussed next. Wenqi et al. [6] proposed an explainable attention-transfer architecture for COVID-19 classification. Their proposed model is implemented as knowledge distillation structures and they evaluated on both CT and chest X-ray image databases. Soumya et al. [7] performed a comprehensive study with eight distinct transfer learning models for COVID-19 detection. They compared the classification outcomes using SqueezeNet, AlexNet, VGG-16, two ResNet variants, Inception, GoogleNet, and MobileNetV2 architectures with chest X-ray image database. They experimented with different parameters such as optimizer type, size of the batch, epochs, and learning rate. Accordingly, they found that ResNet34 architecture provided a maximum of 98.3% classification accuracy. Lucas and Cesar [8] experimented with their work on convolutional neural networks (CNN) for automated identification of COVID severity. Their evaluation is done through chest X-ray image datasets. The maximum performance of 96% was obtained by Inception architecture without any preprocessing approaches. By utilizing a COVID-19 radiographic database, Amit et al. [9] experimented with three distinct transfer learning architectures and averaged the

obtained performance for the binary classification of COVID-19. This binary classification of the disease involves dual cases of COVID-19 negative and positive cases. The research attained a supremacy of 91.6% classification accuracy as compared with other models. Additionally, Maram et al. [10] exposed that the adoption of image augmentation and optimizing the CNN parameters should yield better classification performance as compared with other approaches used for identifying COVID-19. Furthermore, the above approach improved the classification attainment of ResNet50 and VGG-19 architectures for the problem. In addition to this, they proposed CovidXrayNet architecture and attained a maximum classification accuracy of 95.8% with only thirty training epochs. Here, they evaluated their proposed model using two distinct chest X-ray image datasets. Another study [11] proposed a framework for the multiclass problem of COVID-19 detection. Their framework was organized into three phases. First, transfer learning-based ResNet50 was employed for obtaining 2048 feature vectors. Then feature selection was done using Principal Component Analysis (PCA) so that a total of only 64 features were selected. The attributes of the above two steps were then combined and classified finally to attain 98% classification accuracy. Aayush et al. [12] developed a dedicated deep-learning model named SARS-Net to detect COVID-19 severity using chest X-ray image datasets. Here, they employed the COVIDx image dataset with chest X-ray images for experimentation. This leads

their proposed work to obtain a maximum classification accuracy of 97.6% for COVID-19 detection.

In recent days, a type of deep learning called Transfer Learning (TL) has become widely used particularly in the field of biomedical image analysis. By employing five-fold cross-validation and resampling approaches, Ahmad et al. [13] experimented with a study employing a smaller dataset containing 50 positive and 50 negative COVID-19 severities and obtained 98% as maximum accuracy. Luca et al. [14] utilized the power of CNN models for identifying COVID-19 severities with the adoption of three distinct TL networks namely ResNet, VGG, and Xception models. They evaluated their proposed models on a chest X-ray dataset containing pneumonia, normal, and COVID-19 as output classes. Their outcomes revealed that the above-mentioned TL models yielded higher performance whereas the VGG model outperformed other networks. As an extension, the authors proposed further experimentation to boost the TL model's performance by data augmentation and parameter optimization. Wang et al. [15] proposed a TL-based deep-learning architecture called Covid-Net and evaluated it on a database containing approximately 8000 normal, 5530 pneumonia, and 180 COVID-19 cases. In this work, they attained an accuracy of 92% as compared with the existing TL models. Muhammad et al. [16] utilized the widely used ResNet50 model and they concentrated more on fine-tuning the above model to yield better performance. The database that they have employed contains chest X-ray images containing viral pneumonia, bacterial pneumonia, and COVID-19 classes. Shervin et al. [17] adopted a TL-based deep architecture for COVID-19 detection using chest X-ray images. They evaluated this architecture with five thousand X-ray images on the following TL architectures, two ResNet-based models, DenseNet, and SqueezeNet models. Ferhat et al. [18] developed a SqueezeNet-based TL model optimized using the Bayesian algorithm for COVID-19 diagnosis. Due to additional data evaluation and optimal fine-tuning, their work attained improved accuracy and performance. Dian Candra et al. [19] employed a fusion of deep learning and Support Vector Machines (SVM) to classify COVID-19 severities using chest X-ray image databases.

To make the evaluation model more robust, almost all works employing smaller and medium-sized datasets, the images are augmented appropriately and partitioned into training and testing data. In addition, the usage of the attention mechanism has become popular with the intention of enhancing the performance of deep learning architectures. Sivaramakrishnan et al. [20] proposed a methodology involving the augmentation of COVID-19 images with weaker data labels. Asif et al. [21] developed a deep-learning model termed as CoroNet for COVID-19

classification of chest X-ray images. And they adopted a pretrained Xception network evaluated using a four-fold cross-validation strategy. After extracting features, machine learning classification models namely K-nearest neighbor (KNN), SVM, Decision Tree (DT), and Random Forest (RF) models were employed for the COVID-19 task. Rodolfo et al. [22] proposed an approach of stratified analysis for the classification of chest X-ray images corresponding to COVID-19 detection and attained a better performance with an F1 score of 89%. Lal et al. [23] performed COVID-19 classification employing a chest X-ray image dataset with multiclass outputs of Normal, bacterial, and viral pneumonia cases. Chiagoziem et al. [24] developed a CNN-based architecture with second-order pooling, attention mechanism, and dual-path networks for the image analysis of COVID-19 diagnosis. Here, they utilized this dual path for extracting features and second-order pooling for capturing the next-order derivative of the generated feature vectors. And this was done before the implementation of the attention mechanism, thereby attaining better performance outcomes at both the training and testing phases. Recently, Kainat et al. [25] proposed a framework employing nine different TL models for COVID-19 diagnosis utilizing X-ray image datasets. Their results revealed that the VGG-16 model provided maximum performance as compared with other TL architectures. Additionally, the recent studies [26–33] contribute to the classification of COVID-19 severities using neural networks, fine-tuned and optimized transfer learning models for classification using X-ray, CT, and 3D scan images.

From the above discussion, it is apparent that the problem of COVID-19 classification demands reliable and promising methodologies for saving several human lives. Along with the above-discussed reliable methodologies, the work in this paper proposed an improved ResNet50 model i.e., SEA-ResNet50 which includes an attention mechanism, ranger optimizer, and adaptive mish activation function for classifying the chest X-ray images. In this way, the proposed methodology introduced an improved transfer learning approach for tackling overfitting and improving the architecture performance. Hence the research adopted the aforementioned steps to improve the overall performance of the ResNet-50 model substantially.

## Materials and methods

The section provides a detailed discussion of the background of the study, chest X-ray image datasets, pre-processing methods, image augmentation, and proposed transfer learning SEA-ResNet-50 architecture for the employed multiclass classification task.

**Background**

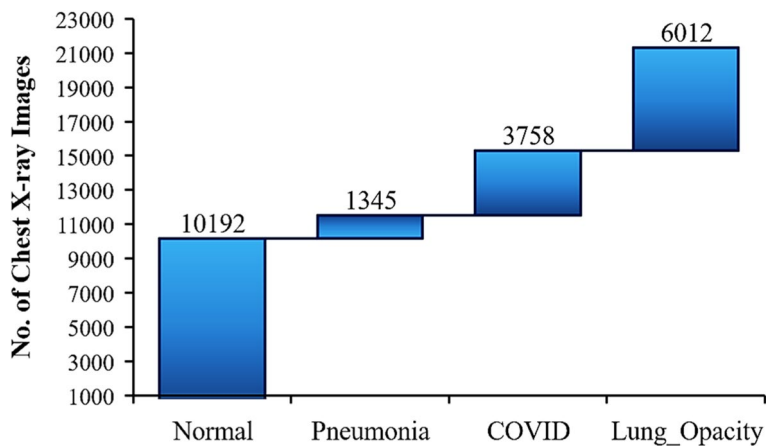
As from the literature discussed in "Related works" section, COVID-19 is a brutal disease all over the world. Thus, there is always a demand for the development of a robust CAD framework for the classification of COVID-19 severity. With the advent of several machine learning (ML) and Deep Learning (DL) algorithms, the research community has been introducing various promising solutions for the severity classification of COVID-19. The background of the study requires a comprehensive understanding of selecting appropriate imaging modality, dataset, preprocessing methods, and classification models.

As follows, several authors as discussed in the previous section have developed promising CAD frameworks. However, the problem still requires a more robust and accurate framework for timely diagnosis. In this way, the work introduces the ResNet50 architecture and attention mechanisms as potential solutions for improving feature representation and classification performance. For COVID-19 severity classification, motivated by the need for a robust CAD framework amidst the pandemic, the study proposes integrating the Squeeze-and-Excitation (SE) mechanism into the ResNet50 model for enhancing its classification performance. This sets the foundation for the research, aiming to design a promising CAD framework for fighting against COVID-19.

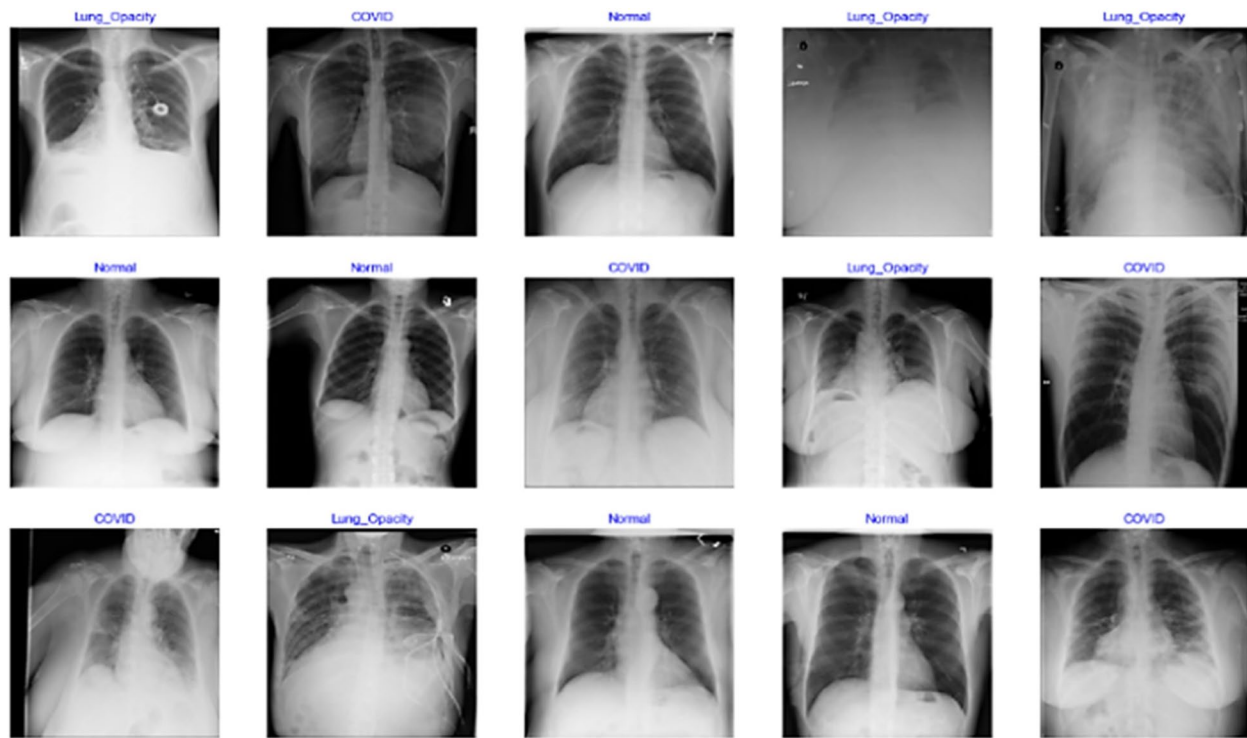
**Chest X-ray image datasets**

The proposed study is evaluated on a combined dataset comprising chest X-ray images attained using two public COVID-19 datasets. This integrated dataset of chest images provides more robustness with lesser prone to variance. The combined dataset contains chest X-ray images with four distinct targets of Normal,

Pneumonia, Lung\_opacity, and COVID-19 cases. The number of X-ray images of the chest available in the two datasets is regularly updated by the dataset developers, and thus the amount of image availability may likely be dynamic in the future. The combined two chest X-ray image repositories are the COVID-19 radiographic dataset [34] publicly available on the Kaggle website and the Chest X-ray radiographic dataset [35] publicly available at the GitHub repository. A research team from the University of Qatar and Dhaka collaborated with medical doctors of Malaysia and Pakistan developed the first dataset containing chest X-ray images with COVID-19-positive, normal, and viral Pneumonia cases. The dataset is composed of 1345 viral Pneumonia, 6012 lung\_opacity, 10,192 healthy (normal), and 3616 COVID-positive chest X-ray images. The second dataset which got approval from the University of Montreal's Ethics Committee contains chest X-ray images of patients who are suspected of COVID-19. Here, the data has been acquired from the public and also involves image acquisition from physicians and hospitals indirectly. From this dataset, 142 COVID-19-positive case chest X-ray images are selected for the research. Thus, the employed chest X-ray images are composed of 10,192 normal, 1345 pneumonia, 6012 lung\_opacity, and 3758 COVID-19 cases. Figure 2 illustrates the chest X-ray data composition of the combined dataset. In addition to this, both the dataset contains chest X-ray images of variable resolution. And it is important to note that the above-selected datasets have no missing values or data for processing. The combined data is then preprocessed for further phases. The sample visualization of chest X-ray images with data labels taken from the combined dataset is given in Fig. 3.



**Fig. 2** Image composition of the combined dataset



**Fig. 3** Chest X-ray samples from the combined dataset

### Preprocessing of chest X-ray images

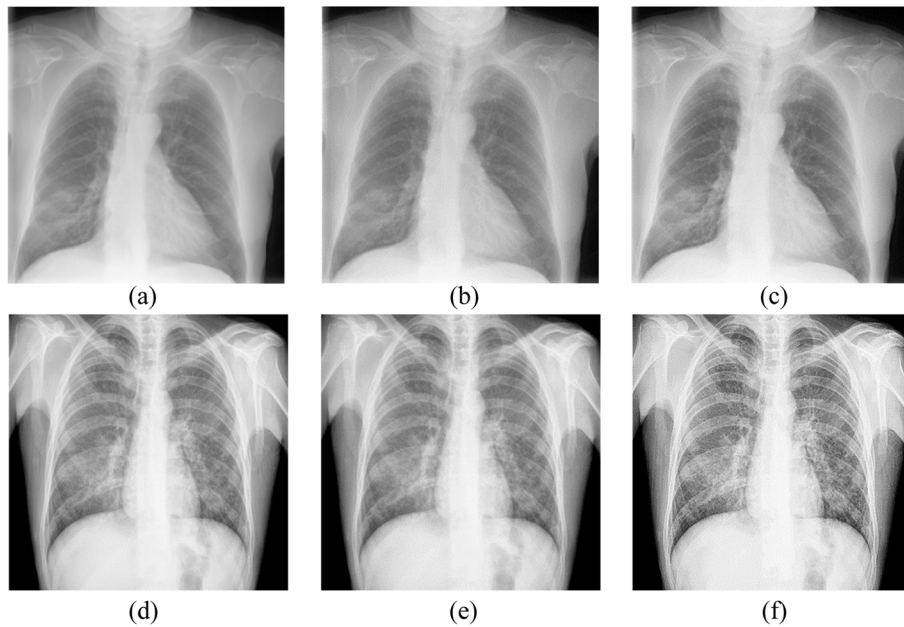
In the design of every CAD model employing medical image analysis, the significance of the preprocessing step is very crucial for attaining better results [36]. The paper employed a simple adaptive median filtering approach [37] for removing any noises included during the acquisition. After filtering, the contrast of the X-ray images is enhanced without any overexposure using Contrast-Limited Adaptive Histogram Equalization (CLAHE) [38]. This is performed for the grayscale equalization of the chest X-ray images. Figure 4 illustrates the chest X-ray images of each stage of preprocessing where 4(a) denotes the raw image with the COVID label, 4 (b) represents the noise-removed image using an Adaptive Median Filter (AMF), and 4 (c) indicates the Contrast-Limited Adaptive Histogram Equalization (CLAHE) image. The histogram visualization of the adaptive median filtered and the CLAHE processed images are illustrated in Fig. 5. The steps involved in CLAHE processing are given below. Let  $H(i)$  be the histogram of pixel intensity values in the input image, where  $i$  ranges from 0 to  $L - 1$ . Here,  $L$  indicates the amount of intensity levels of the input image. The cumulative distributed function (CDF) is then computed as the cumulative sum of the histogram [39] as given in Eq. (1).

$$CDF(i) = \sum_{j=0,1 \dots i} H(j) \quad (1)$$

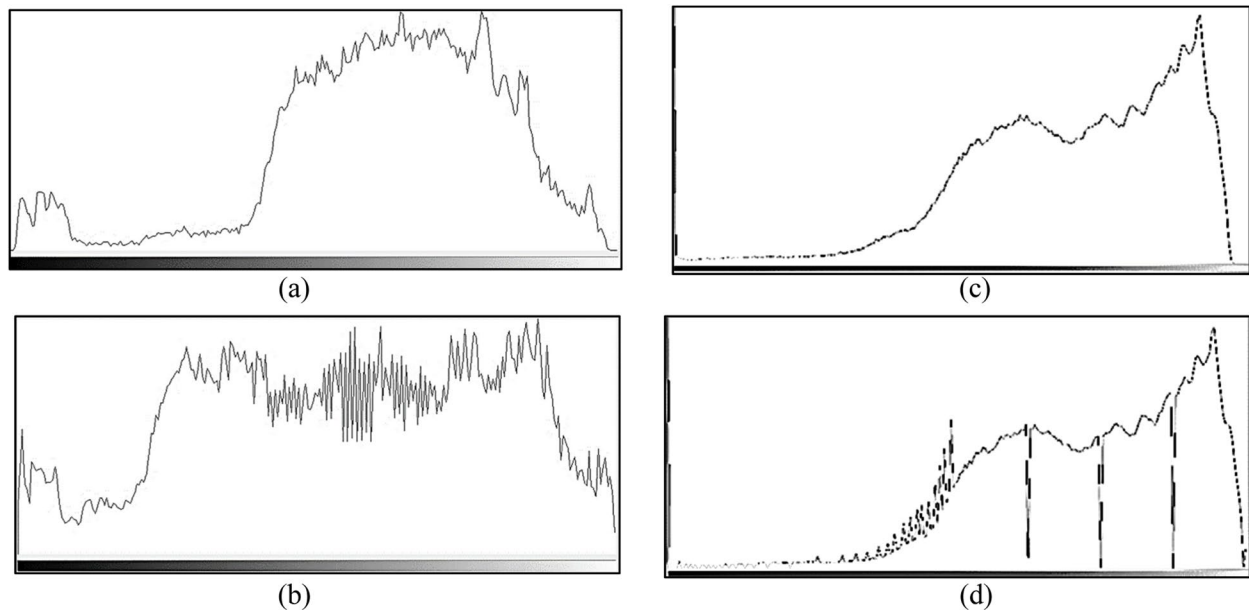
The transformation function  $T(i)$  maps the original intensity values to the enhanced intensity [39] as given in Eq. (2).

$$T(i) = \text{round} \left( \frac{CDF(i) - CDF_{min}}{(M \times N) - CDF_{min}} \right) \times (L - 1) \quad (2)$$

In Eq. (2),  $CDF_{min}$  indicates the minimum CDF value in the local neighborhood,  $(M \times N)$  represents the size of the local neighborhood, and the number of intensity levels is indicated as  $L$ . The work employs the value of 16 for both  $M$  and  $N$ . The Eqs. (1) and (2) illustrate the process of computing the cumulative distribution function and the transformation function used in CLAHE. This provides the enhancement of local contrast of chest X-ray images, which is vital for accurate classification of different medical conditions in a multiclass classification problem. The next step of CLAHE involves clipping the transformation function to limit the amplification of noise as shown in Fig. 5. After CLAHE transformation, the chest X-ray images of the combined dataset are augmented for increasing the amount of data and for robust evaluation of the



**Fig. 4** Preprocessing of Chest X-ray data **(a, d)** original images (COVID label) of datasets 1 and 2 **(b, e)** adaptive median filtered images of datasets 1 and 2 **(c, f)** CLAHE images of datasets 1 and 2



**Fig. 5** Histogram of a sample chest X-ray Image, **(a, c)**—adaptive median filtered output of a sample image of datasets 1 and 2; **(b, d)**—CLAHE output of a sample image of datasets 1 and 2

research work. This image augmentation is carried out by varying the zoom and shear parameters of the processed inputs. In this way, the work made five copies

of each processed chest X-ray image. Then, the augmented data are applied to the different transfer learning models.

### Transfer learning models for multiclass classification of COVID-19 chest X-ray images

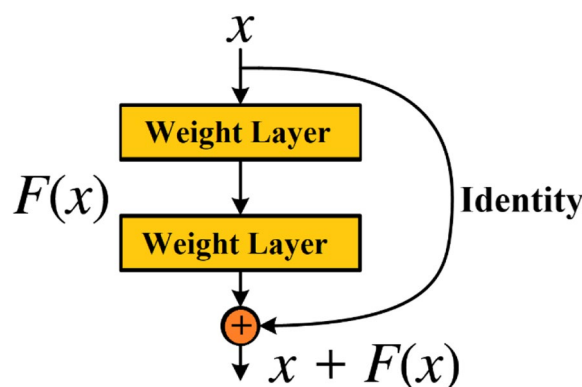
The details on different transfer learning models employed for multiclass classification of chest X-ray image data are discussed in this section. After cropping the processed images to  $224 \times 224 \times 3$ , they are fed to the deep learning models as discussed below. These transfer learning models exist with already pretrained on the ImageNet database [40]. This database is made of 14,197,122 generic image data with around a thousand object classes. This makes this dataset benchmark data for detecting objects, visual recognition, and image classification [40].

#### Transfer learning model: VGG-16

Due to the ability to learn discriminative features from the input images, the VGG-16 model is employed for the COVID-19 classification. The architecture consists of 13 convolutional layers followed by 3 fully connected layers [41]. Here, the convolutional layers are organized in blocks with each one containing multiple convolution layers followed by a max-pooling layer. Accordingly, the pre-processed images are passed through the convolutional layers, applying activation functions (ReLU), and down-sampling through max-pooling layers [41]. After the convolutional layers, the feature maps are then flattened into a vector and applied through fully connected layers. Those attained outputs are then applied through a softmax activation function for obtaining class probabilities [41].

#### Transfer learning model: Xception

The proposed work employs the Xception transfer learning model due to its excellence in better feature extraction and representation. The Xception architecture is actually an extension of the Inception model and is well-known for its unique feature i.e., depth-wise separable convolutions [42]. This uniqueness of the model helps in reducing computational complexity while preserving representational capacity. The architecture consists of a series of depth-wise separable convolutional layers, followed by global average pooling (GAP) and fully connected (FC) layers [42]. Herein, the depth-wise separable convolutions should be used to split the standard convolution operation into two separate operations: the first one is depth-wise convolution and the next one is point-wise convolution. This point-wise convolutional operation helps the model to have reduced computational complexity [42].



**Fig. 6** Concept of residual model (skip connection)

#### Transfer learning models: ResNet18 and ResNet50

In conventional CNNs, each layer is expected to learn a direct mapping from inputs to outputs. The concept of residual learning introduces a shortcut connection as illustrated in Fig. 6. These shortcut connections are also termed as skip connections in the Residual transfer learning model. The skip connection as illustrated in Fig. 6 bypasses one or more layers, enabling the residual architecture to learn residual functions [43]. Thus, the architecture helps in the problem of vanishing gradient problem in deep learning models. The output of the residual block  $H(x)$  can be calculated [43] as shown in Eq. 3.

$$H(x) = F(x) + x \quad (3)$$

In Eq. (3),  $x$  portrays the residual block's input, and  $F(x)$  indicates the residual function to be learned within the block. This equation represents the core idea of a residual block. And this signifies the concept that by adding the input  $x$  directly to the output  $F(x)$ , the original information is preserved and propagated through the network. As a result, this helps in training stability and accuracy.

As shown in Fig. 6, the basic residual block structure contains two convolutional layers with batch normalization and ReLU activation functions. In the first convolutional layer, the convolution operation is done followed by batch normalization and ReLU activation function. Let  $W_1$  as the weight matrix corresponds to the first convolution layer, and  $b_1$  as the bias vector. This makes the output ( $O_{First}$ ) of the first convolution layer [44] as given in Eq. (4).

$$O_{First} = ReLU(W_1 * x + b_1) \quad (4)$$



In Eq. (4), the ReLU (Rectified Linear Unit) activation function is applied to introduce non-linearity and this helps the network to learn complex patterns. In the second convolution layer, another convolution operation is performed followed by batch normalization. Here, consider  $W_2$  as the weight matrix corresponding to the second convolution layer, and  $b_2$  as the bias vector. This makes the output ( $O_{Second}$ ) of the second convolution layer [44] as given in Eq. (5).

$$O_{Second} = W_2 * ReLU(W_1 * x + b_1) + b_2 \tag{5}$$

The Eqs. (4) and (5) state that batch normalization is typically applied after the convolution and before the ReLU activation to normalize the activations and improve training. In this way, the input  $x$  will be summed with the outputs of the second convolution layer while preserving the original information. Thus the final output of the residual block [44] can be given as represented in Eq. (6).

$$H(x) = W_2 * ReLU(W_1 * x + b_1) + b_2 + x \tag{6}$$

Equation (6) states that by adding  $x$ , the original input information is preserved and helps in gradient flow which enables the training of deeper networks. ResNet18 and ResNet50 consist of several residual blocks, each containing multiple convolutional layers. Here, ResNet18 consists of eighteen layers (18 weight layers) and ResNet50 consists of fifty layers [45]. Thus, both ResNet18 and ResNet50 architectures comprise convolutional layers, batch normalization, ReLU activations, and skip connections. The architectural diagram of the ResNet50 transfer learning model is illustrated in Fig. 7. This diagram shows that the ResNet50 model has five stages with 50 deep layers. As shown, it is a powerful CNN architecture used for addressing the challenges of training very deeper networks. This could be achieved by introducing residual connections as shown in Figs. 6 and 7. The model seems to be effective in learning intricate patterns and features from any complex datasets. Due to these advantages, the

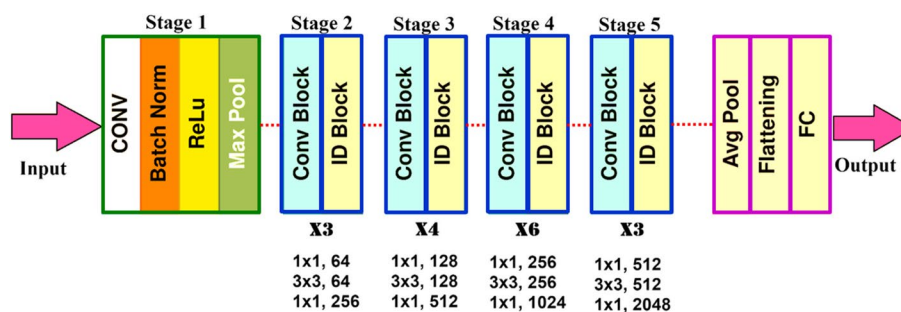
model has become popular in several image recognition problems [46]. Finally, the residual mapping  $F(x)$  can be expressed [46] as given in Eq. (7).

$$H(x) = W_3 * ReLU(W_2 * ReLU(W_1 * x + b_1) + b_2) + b_3 \tag{7}$$

In Eq. (7),  $W_1$ ,  $W_2$ , and  $W_3$  represent the weight matrices of the convolutional layers within the blocks,  $b_1$ ,  $b_2$ , and  $b_3$  denote the respective bias vectors, the symbol  $*$  indicates the convolution operation, and  $ReLU$  represents the rectified linear unit activation function. The aforementioned equations discussed in this sub-section illustrate the flexibility and capability of ResNet architectures to model intricate patterns and features from complex datasets. Several recent studies [47–50] in medical image analysis have demonstrated the superior performance of ResNet50 compared to other architectures and the model is noteworthy for biomedical image analysis due to its balance between depth and efficiency.

**Transfer learning model: DenseNet121**

This type of pretrained architecture is well-known for its deep convolution neural network with densely connected layers. That is, DenseNet stands for densely connected neural networks in which they introduce dense connectivity patterns between layers. This provides the advantages of feature reuse and gradient flow throughout the network [51]. The architecture contains several dense blocks, each containing multiple convolution layers. This implies that within a dense block, each layer is interconnected to all others in a feed-forward fashion. And the outcome of each layer should be concatenated with the inputs of all subsequent layers within the similar block [51]. In addition, transition layers are included to control the spatial dimensions of feature maps between dense blocks. Between dense blocks, these layers are utilized for down-sampling feature maps. This is done by exercising convolution and average pooling operations. This will help to reduce the spatial dimensions of feature maps and control model complexity [51].



**Fig. 7** The architecture of ResNet50 transfer learning model

### Proposed Squeeze-and-Excitation Attention—based ResNet50 (SEA-ResNet50) model for COVID-19 detection

The proposed work intends to improve the ResNet50 transfer learning architecture for improving the performance of multiclass COVID-19 classification utilizing chest X-ray data. To attain the efficient utilization of feature vectors across the channels according to their significance, the Squeeze-and-Excitation Attention (SEA) [52] model is introduced. At first, this SEA unit has to be integrated with the residual units, and the implementation is similar to the simple attention concept. Additionally, the adaptive Mish activation function is introduced into the architecture. This step is required in the proposed work for avoiding neuron necrosis occurring due to the forced sparse mechanism of the rectified linear unit activation function. Consequently, an effective Ranger optimizer is taken for further improving the performance of the ResNet50 architecture.

#### Why Squeeze-and-Excitation Attention (SEA) unit?

In deep CNN models, several research works are intended to improve their performance through the stacking of convolution layers. This is done to concatenate more special feature maps for the considered problem. However, this will increase the depth of models resulting in complex training. Additionally, the existing CNNs have the rare ability to discriminate the significance of each channel while performing feature concatenation. This will result in under-appreciation of some significant feature channels and hence the overall performance will be reduced. To overcome these issues, the Squeeze-and-Excitation Attention (SEA) module is employed for fusing features from the channel dimensions. This is a useful imitation of biological attention mechanisms that have the ability to concentrate highly on informative channel features and neglect the least significant features for enhancing classification performance.

#### Squeeze-and-Excitation Attention (SEA) unit

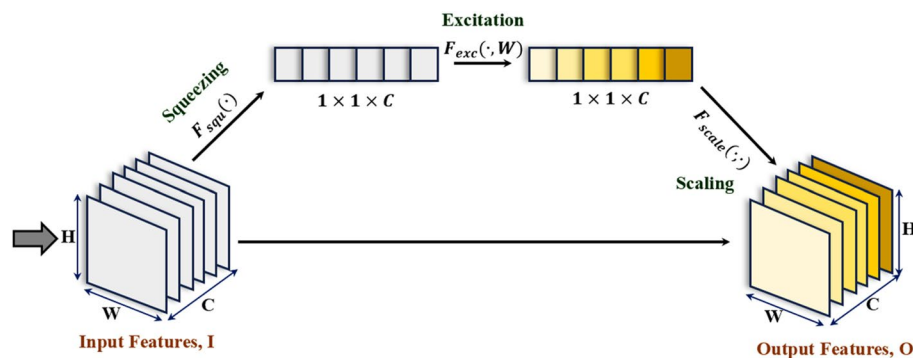
This type of attention unit belongs to an attention mechanism that can be integrated into deep CNNs such as the ResNet50 model. This is done to enhance feature representations and improve model performance, especially for the employed multiclass classification task of COVID-19 detection using chest X-ray images. The representation of the squeeze-and-excitation attention (SEA) module used in the residual units of the work is illustrated in Fig. 8. As shown, this module consists of only two operations namely squeeze and excitation operations. This starts with the compression of the input features,  $I$  in the spatial dimension for each channel,  $C$ . And this operation is termed as squeezing operation,  $F_{squ}$ . Afterward, in excitation operation,  $F_{exc}$ , these compressed vectors are passed to two FC layers. Accordingly, the weights of each feature channel will be created using the Sigmoid equation. Then, the original feature maps are multiplied (scaling operation) by the channel-wise weights learned during the excitation operation. In this way, the proposed model focuses more on informative channels and suppresses less relevant ones. This will lead the work to attain enhanced feature representations for COVID-19 classification utilizing chest X-ray data. Thus, the SEA module is utilized to capture channel-wise dependencies and for enhancing feature representation adaptively. This makes the architecture to focus on informative regions of the input feature maps while suppressing non-significant ones. The mathematical representation can be given as follows.

$$Z_c = F_{squ}(u_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (8)$$

$$s = F_{exc}(Z, W) = \sigma(W_2 \sigma(W_1 Z)) \quad (9)$$

$$O = F_{scale}(u_c, s_c) = s_c \cdot u_c \quad (10)$$

In Eq. (8),  $Z_c$  and  $\sigma$  represent the values of the feature vectors corresponding to distinct  $c_{th}$  channel



**Fig. 8** Representation of Squeeze-and-Excitation Attention (SEA) module

and sigmoid function.  $F_{squ}$ ,  $F_{exc}$ , and  $F_{scale}$  denote the operations of squeezing, excitation, and scaling,  $u_c(i, j)$  denotes the eigenvalue at  $(i, j)$  of  $c^{th}$  channel, and  $W_1$ ,  $W_2$  indicate the weights corresponding to two FC layers. Here, Eq. (8) signifies the way of obtaining a single scalar value ( $Z_c$ ) for each channel through squeeze operation. As illustrated, this operation condenses the spatial information of each channel into a single value and provides the ability to capture the global distribution of features in that channel. Conversely, this results in the reduction of spatial dimensions while retaining crucial channel-wise information. Equation (9) signifies the concept of excitation operation that allows the model to learn and emphasize the importance of different channels dynamically, based on the global context. Finally, Eq. (10) illustrates the scaling operation which signifies that by individual scaling of each channel, the model can enhance or suppress certain features. This will improve the model's ability to focus on relevant information and ignore irrelevant details. On the whole, the aforementioned equations illustrate the operations of the SEA unit which highlights its role in recalibrating channel-wise feature responses through squeezing, excitation, and scaling operations.

#### Adaptive mish activation function

In convolutional neural networks, the activation functions play a significant role in deciding the activation and deactivation of neurons. This is done by computing the weighted sum of inputs and biases in a network. As compared with the standard non-linear activation functions such as ReLU, Leaky ReLU, and Swish, the Mish activation function aims to provide smoother gradients and better generalization performance for classification tasks [53]. It is mathematically represented [53] as given in Eq. (11).

$$Mish(x) = x \cdot \tanh(\ln(1 + e^x)) \quad (11)$$

The exponential term of Eq. (11) ensures that for larger positive inputs, the function can grow rapidly. The logarithmic term moderates the exponential growth and helps in making the function remains positive. The hyperbolic tangent term in the equation helps to squashes the output and thus adding smooth non-linearity. Hence, the mathematical formulation of the Mish activation function combines the aforementioned functions to achieve a smoother, non-monotonic alternative to traditional

activation functions. This, in turn, provides benefits such as better gradient flow, improved generalization, and the ability to capture complex patterns. This makes it a powerful tool for improving the performance of convolutional neural networks in classification tasks and beyond.

#### Adaptive parameter ( $\alpha$ )

The proposed work introduced an adaptive Mish activation function as represented in Eq. (12). The Adaptive Mish activation function is introduced to adaptively adjust in accordance with the characteristics of input data. The idea is to introduce learnable parameters as shown in Eq. (12). This is used in such a way that dynamically modulates the Mish function's response so that this setup will capture the intricate features present in COVID-19 chest X-ray images.

$$AdaptiveMish(x) = x \cdot \tanh[\alpha \cdot (\ln(1 + e^x))] \quad (12)$$

In Eq. (12),  $\alpha$  represents the introduced learnable parameter. This  $\alpha$  is used to control the adaptive behavior of the activation function whereas  $\tanh$  denotes the hyperbolic tangent function. As in Eq. (12),  $\alpha$  allows the Mish activation to adapt its curvature and slope based on the applied input images. Here, the value of  $\alpha$  has been adjusted adaptively based on the learning rate during training epochs. And during backpropagation, the gradients of the loss function with respect to  $\alpha$  are computed. These gradients guide the updates to  $\alpha$  using a gradient descent-based optimization algorithm. This makes the activation function to have two significant advantages. The first one is able to capture varying degrees of non-linearity effectively and the next one is ensuring smoother gradients during training. Within the SEA modules, the adaptive parameter  $\alpha$  can be learned alongside other parameters. This enables the SEA-ResNet50 architecture to dynamically adjust attention weights based on the input features. This adaptability will lead to obtain enhanced model performance and improved generalization capacity and thus influence more accurate multiclass classification results for the employed problem. Figure 9 portrays the comparison of Mish [53] and Adaptive Mish activation functions for different alpha values. This plot illustrates the adaptability of the Adaptive Mish function, which can be tuned adaptively based on the applied chest X-ray images. The final overall structure of the SEA-ResNet50 with modified Mish activation function is illustrated in Fig. 10. The algorithmic summary of the proposed model is given below.

**Algorithm 1.** SEA-ResNet50 with adaptive Mish activation and ranger optimizer for COVID-19 detection

*Input: Chest X-ray Images from the Combined Dataset*

**Step 1:** Load chest X-ray images with a batch size of 16.

**Step 2:** Noise removal using adaptive median filtering approach.

**Step 3:** Contrast Limited Adaptive Histogram Equalization (CLAHE) based contrast enhancement using Equations (1) and (2).

**Step 4:** Image Augmentation and resizing of preprocessed inputs to the fixed size of  $224 \times 224 \times 3$ .

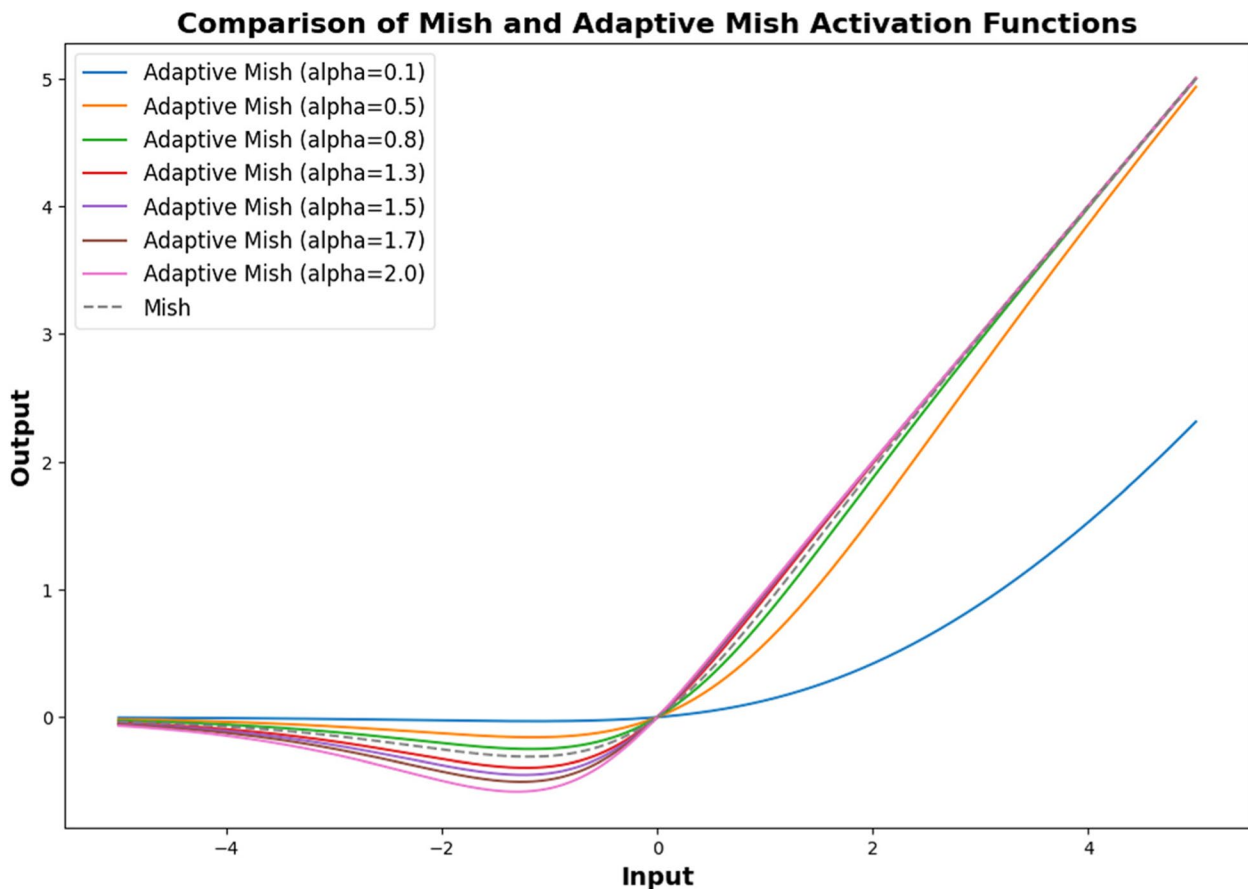
**Step 5:** Integration of the Squeeze-and-Excitation visual attention units to the residual blocks of ResNet50 architecture as mentioned in Equations (7) to (10).

**Step 6:** Replacing the ReLU with the adaptive Mish activation function as mentioned in Equation (12).

**Step 7:** Stratified partition of augmented image inputs to 70:30 ratio of training and testing data. The work follows 70% for training and validation with 5-fold cross-validation, while 30% is for testing.

**Step 8:** Employment of five TL models (VGG-16, Xception, ResNet18, ResNet50, and DenseNet121) for comparative study with the proposed SEA-ResNet50 architecture.

**Step 9:** Training and Testing evaluation of the aforementioned deep learning models using Ranger Optimizer.



**Fig. 9** Comparison of Mish and adaptive Mish activation function for different alpha values

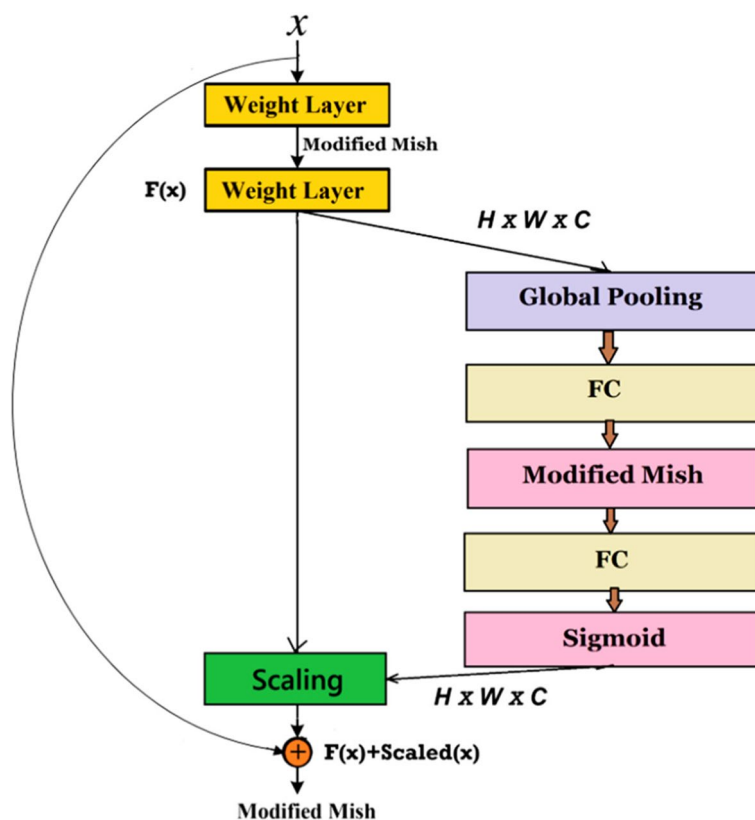


Fig. 10 Overall structure of the SEA-ResNet50 with adaptive Mish activation function

### Experimental outcomes and analysis

The analysis of the attained results utilizing the proposed approach for the multiclass COVID-19 classification has been portrayed here. The implementation and evaluation of the proposed work are executed with Jupyter Notebook on a recent version of the Windows operating system having 16 GB RAM. For comparative study, the paper utilized the five existing TL models namely VGG-16, Xception, ResNet18, ResNet50, and DenseNet121 architectures. The evaluation of the study is done using a partition ratio of 70% training and 30% testing data of preprocessed X-ray images. The above-mentioned ratio follows the stratified partition of splitting data. To tackle the problem of class imbalance, a fivefold cross-validation is applied within the 70% training data for tuning, and the final model is evaluated on the 30% testing set. Based on the experimental validation, the learning of the model is considered to be  $1.0 \times 10^{-3}$  with 16 as batch size, optimizer as Ranger (combining RAdam and Lookahead), activation function as Adaptive Mish, and 20 as number of epochs. The above hyperparameters are chosen carefully through a combination of grid searches, preliminary experiments, and benchmarking. Thus, the paper ensured the optimal performance of the SEA-ResNet50

model in detecting COVID-19 from chest X-ray images. To mitigate the class imbalance problem, data augmentation, attention mechanism, cross-validation, and robust performance metrics such as precision, F1 score, and Cohen’s kappa score are employed.

### Impact of individual modules used to improve the ResNet50 architecture

In this sub-section, 5 individual ablation tests are done for further exploring the effectiveness of the proposed methodology. This tests the impact of the employed SEA units, modified Mish activation function, and Ranger optimizer for the employed task using preprocessed COVID-19 chest X-ray images. Table 1 summarizes the outcomes of these ablation tests for COVID-19 detection using ResNet50 architecture.

At first, in Test\_1, the ResNet50 model is evaluated without any changes using the image inputs, and 94.31% classification accuracy is obtained. In Test\_2, the residual units in the standalone ResNet50 architecture are replaced with the residual blocks enhanced using the SEA block. This is done to ensure the effectiveness of the SEA unit in ResNet50 for this problem. As in Table 1, the test provided an improved accuracy of 2.22% for using

**Table 1** Ablation tests to know the impactness of the individual modules used in the resnet-50 architecture

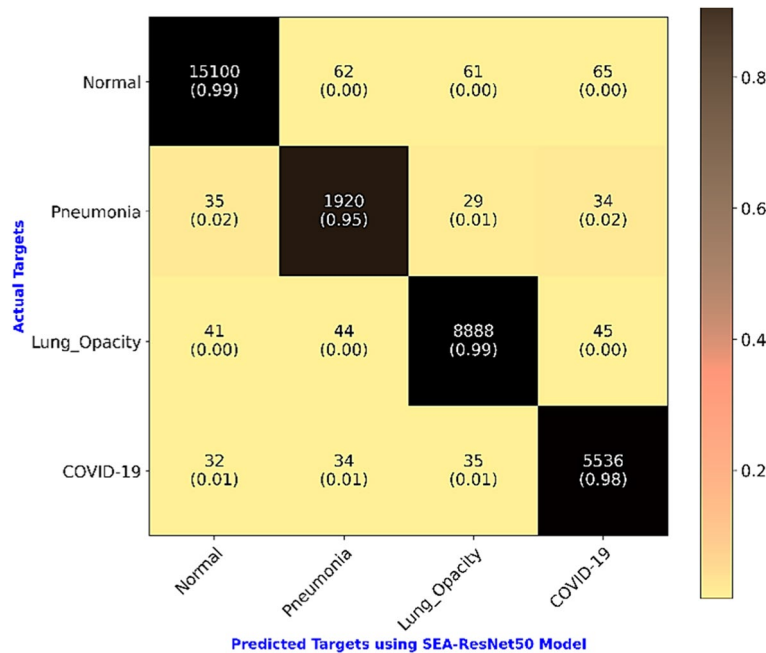
Modules	Test_1	Test_2	Test_3	Test_4	Test_5
SEA Unit	✗	✓	✗	✓	✓
Adaptive Mish Activation	✗	✗	✓	✓	✓
Ranger Optimizer	✗	✗	✗	✗	✓
<b>Accuracy (%)</b>	<b>94.31</b>	<b>96.53</b>	<b>95.12</b>	<b>97.95</b>	<b>98.38</b>

only the SEA module in ResNet50. This test result indicates that the employment of this SEA block for distinguishing the significance between the channel features has an impact on building a better classification model for COVID-19 detection. In Test\_3, the ReLU activation functions are replaced with the adaptive Mish activation function. This enhanced the classification accuracy improvement of 0.81% over the standalone ResNet50. This implies that the replaced activation provides a strong regularization effect for chest X-ray data classification. In Test\_4, the combination of SEA block and adaptive Mish activation function is utilized in the ResNet50 model and so an accuracy improvement of 3.64% is obtained. In Test\_5, the most commonly employed Adam optimizer is

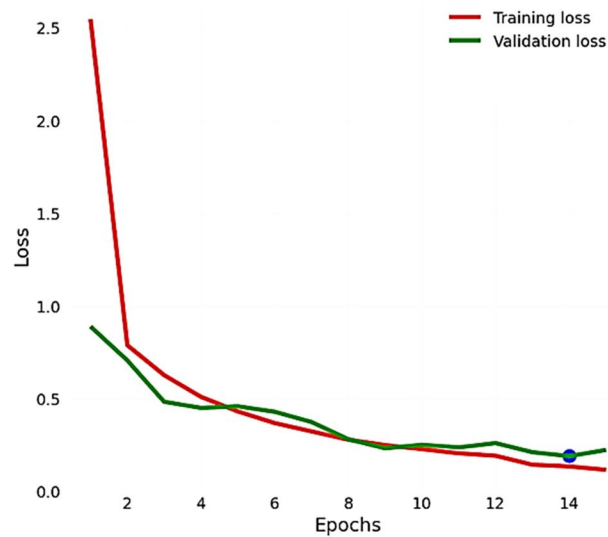
replaced with the Ranger optimizer [54] in the ResNet50 model. This provides an accuracy improvement of 0.43% over the previous test. This is due to the fact that the Ranger is simply the combination of Rectified Adam and Lookahead [54]. The above ablation test results indicate that the combination of the SEA model, adaptive Mish activation function, and the Ranger optimizer is found to be reliable and effective for COVID-19 chest X-ray image classification.

**Multiclass classification: outcomes of the proposed approach and its comparative analysis**

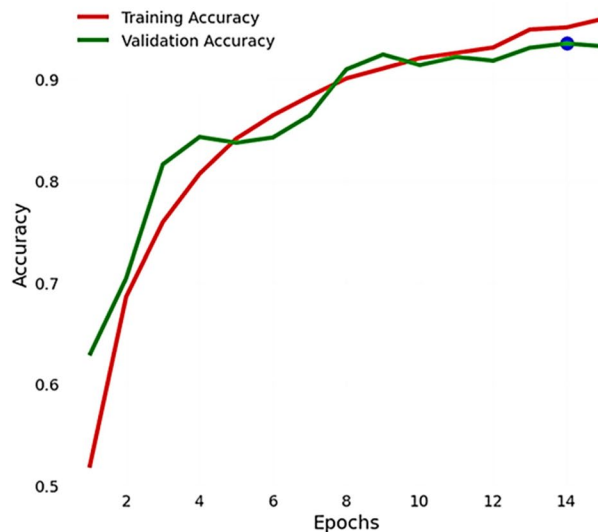
After the implementation of the SEA\_ResNet50 model for the considered problem, the confusion matrix is



**Fig. 11** Confusion matrix obtained for multiclass classification using SEA-ResNet50 model



(a) Training and Validation Loss Plot of SEA-ResNet50 for Multiclass COVID-19 Classification Problem



(b) Training and Validation Accuracy Plot of SEA-ResNet50 for Multiclass COVID-19 Classification Problem

**Fig. 12** Training and validation plots of SEA-ResNet50 for multiclass COVID-19 classification problem

obtained. To attain these results, a five-fold cross-validation strategy is employed. From this matrix corresponding to four classes namely Normal (Class\_1), Pneumonia (Class\_2), Lung\_Opacity (Class\_3), and COVID-19 (Class\_4) cases, the performance metrics are calculated. This includes recall, precision, accuracy, F1-score, macro-F1, and weighted-F1 measures [55]. And these attained outcomes are then validated using Cohen's kappa ( $\kappa$ ) validation score [56]. Figure 11 illustrates the confusion matrix for the employed multiclass classification problem using the SEA-ResNet50 architecture.

Figure 12 illustrates the validation and training plots of the SEA-ResNet50 model applied for multiclass COVID-19 detection. As shown in this plot, the proposed model provided supreme performance at epoch 14. The attained results in the testing phase for this four-class classification of COVID-19 severity are tabulated in Table 2. Here, Macro-F1 represents the average F1 score across all four classes so that equal weightage to each class is given. And Weighted-F1 represents the weighted average of the F1-score across all classes, considering the class imbalance. This metric accounts for

**Table 2** Classifier's performance for multiclass COVID-19 detection

DL models	Precision (%)				Recall (%)				F1-Score (%)				Overall accuracy (%)	Macro-F1 (%)	Weighted-F1 (%)
	Class_1	Class_2	Class_3	Class_4	Class_1	Class_2	Class_3	Class_4	Class_1	Class_2	Class_3	Class_4			
	VGG-16	92.62	80.67	91.66	90.05	96.16	66.72	92.28	86.95	94.36	73.04	91.97			
Xception	93.73	83.65	92.75	90.93	96.63	70.45	93.31	88.81	95.16	76.48	93.03	89.86	92.33	88.63	92.21
ResNet18	95.04	85.63	93.88	91.64	97.09	74.42	94.26	90.76	96.06	79.63	94.07	91.20	93.52	90.24	93.44
ResNet-50	95.70	86.62	94.36	93.24	97.43	77.14	95.14	91.65	96.55	81.61	94.74	92.44	94.31	91.34	94.25
DenseNet121	96.55	88.11	95.02	94.84	97.85	80.71	96.02	92.97	97.19	84.25	95.52	93.90	95.28	92.71	95.24
Proposed SEA-ResNet50	98.77	95.14	98.56	98.21	99.29	93.20	98.61	97.46	99.03	94.16	98.59	97.84	98.38	97.40	98.36



the number of samples in each class and offers a more accurate measure of DL architecture performance. As listed in Table 2, the VGG-16 model provided a macro F1 score of 86.96%, an overall accuracy of 91.14%, and a weighted F1 score of 90.99%. However, the Xception model provided a better classification performance than VGG16 for multiclass COVID-19 classification. That is, it provided a macro F1 score of 88.63%, an overall accuracy of 92.33%, and a weighted F1 score of 92.21%.

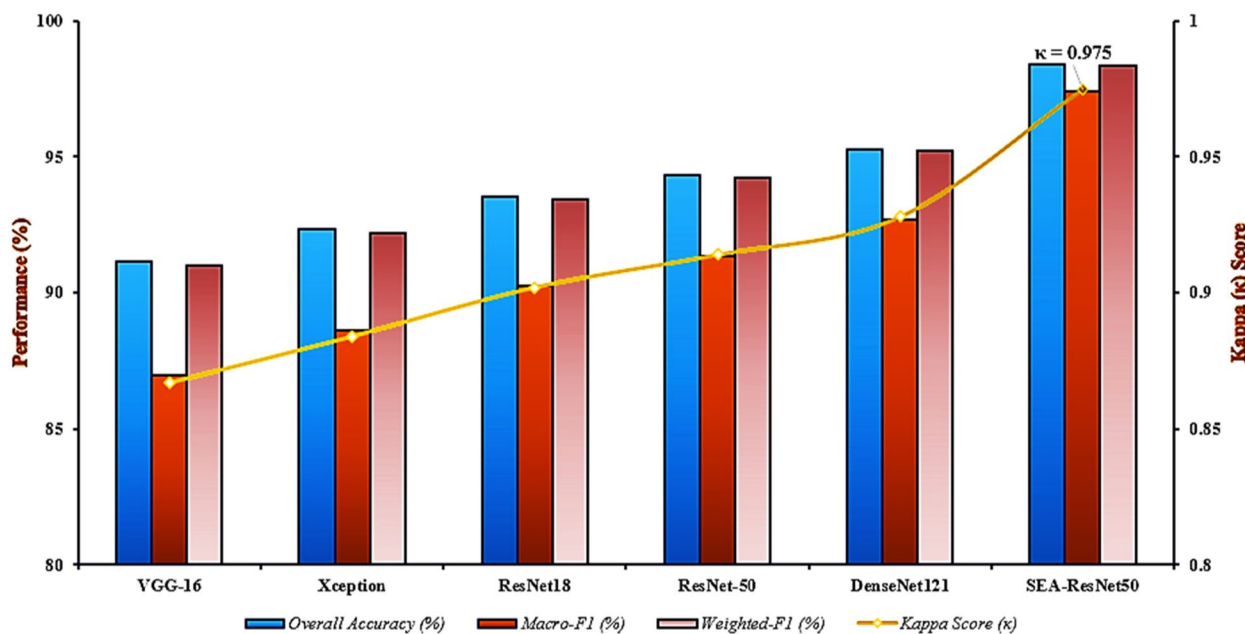
The classification performance has been increased further when skip connections-based deep learning models are employed. This reveals that the skip connections used in the ResNet18 and ResNet50 models make a significant contribution to the employed problem. In this way, the ResNet50 model provided a better classification performance of 94.31% overall accuracy, macro F1 score of 91.34%, and weighted F1 score of 94.25%. And also noted that the class-wise performance is substantially improved over the aforementioned transfer learning models.

Afterward, the densely connected DenseNet121 architecture was utilized but found that the results of multiclass classification for the employed problem notably overlapped with the previous ResNet50 model. So, the performance of the ResNet50 model is further improved using the SEA blocks, modified mish activation, and Ranger optimizer. Thus, the proposed SEA-ResNet50 architecture provided a macro F1 score of 97.40%, an overall accuracy of 98.38%, and a weighted

F1 score of 98.36%. Also, the individual class-wise performance is substantially improved over other models. Consequently, the proposed SEA-ResNet50 model's superior performance could be attributed to its integration of the Squeeze-and-Excitation attention mechanism, effective feature representation capabilities, usage of ranger optimizer, and experimental setup. The above-discussed factors contributed collectively to the model's ability to attain higher precision, accuracy, recall, and F1 scores in the problem of multiclass COVID-19 detection. The above performances of DL models are validated using Cohen's score and plotted graphically in Fig. 13. The plot validates the superior performance of the proposed model with a kappa score of 0.975.

**Binary classification: outcomes of the proposed model and its comparative analysis**

The results obtained for binary classification of COVID-19 detection, i.e., considering only two cases, Normal (Class\_1) and COVID-19 (Class\_2) are summarized in Table 3. Herein, it is revealed that the skip connection-based ResNet models provided better classification performance for the binary class problem. That is, ResNet50 provided a better classification performance of 94.91% sensitivity, 96.60% specificity, 96.14% accuracy, 91.14% precision, and 92.99% F1-score. After introducing the squeezing-and-excitation attention blocks, modified mish activation, and Ranger optimizer, the SEA-ResNet50 model outperformed others with 99.13%



**Fig. 13** Comparative plot analysis of transfer learning models for multiclass classification of COVID-19 using kappa validation

**Table 3** Binary classification outcomes of the proposed approach for COVID-19 diagnosis

DL models	Sensitivity (%)	Specificity (%)	Accuracy (%)	Precision (%)	F1 Score (%)
VGG-16	88.88	93.83	92.49	84.15	86.45
Xception	91.91	94.52	93.82	86.08	88.90
ResNet18	94.00	95.29	94.94	88.04	90.92
ResNet-50	94.91	96.60	96.14	91.14	92.99
DenseNet121	95.80	97.38	96.96	93.10	94.43
Proposed SEA-ResNet50	99.13	99.35	99.29	98.24	98.68

of sensitivity, 99.35% of specificity, 99.29% of accuracy, 98.24% of precision, and 98.68% of F1-score values. Figure 14 illustrates the comparison of accuracy and kappa scores for binary classification of COVID-19 detection. As from the plot, the proposed work attained the supreme kappa validation score of 0.98 as compared with the existing models.

#### Saliency map visualization and interpretation

The visualization of salient class activation maps influenced for COVID-19 multiclass classification using the proposed model is illustrated in Fig. 15. This visualization is done using Grad-CAM-based [57] Explainable AI which reveals that the proposed SEA-ResNet50 provides supreme classification performance over others. The steps involved are given as follows.

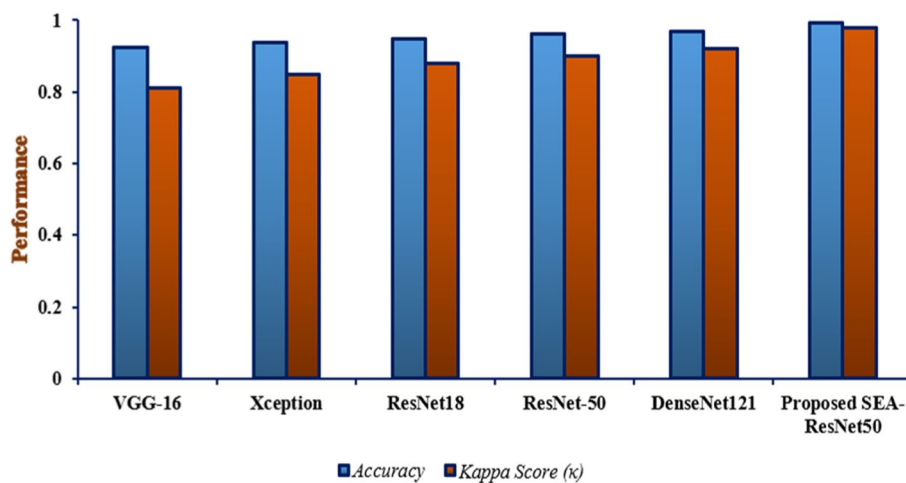
- (i) Gradient Calculation: Grad-CAM computes the gradient of the score for a target class  $y^c$  (e.g., pneumonia, COVID-19) for an applied input chest X-ray image. This calculation is done by considering the feature maps  $A^k$  of the final convolutional layer. The resultant gradients ( $\partial y^c / \partial A^k$ ) will signify how important each feature map is for the output target.

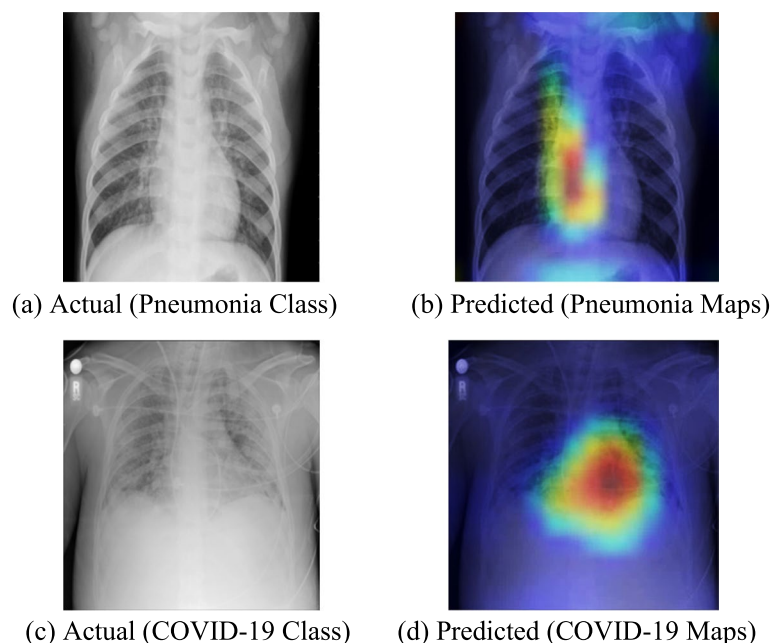
- (ii) Weighted Feature Maps: The above-resultant gradients are then averaged (global average pooling) for obtaining the weights  $\alpha_k^c$  [58] as given below.

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A^k} \quad (13)$$

Equation (13) characterizes the importance of each feature map  $k$  for the target class  $c$  and  $Z$  denotes the number of pixels in the feature map. Equation (13) provides a way to quantify the importance of each feature map in a convolutional neural network for a specific target class. By averaging the gradients of the class score for the feature map activations, it captures the overall influence of the feature map on the class prediction. This information is essential for understanding, interpreting, and improving the performance of deep learning models, particularly in complex tasks like image classification.

- (iii) Class Activation Map: The above weights are then utilized for performing a weighted combination of the feature maps. Next, a ReLU activation function is employed to make sure to consider only the positive contributions. As a result, a coarse heatmap is

**Fig. 14** Accuracy and kappa score comparison for binary classification of COVID-19 detection



**Fig. 15** Visualization of class activation map using the proposed model

obtained that highlights the regions of the image that are most relevant for the SEA-ResNet50 model's prediction.

- (iv) **Overlaying Heatmaps:** With respect to the input image, the above heatmaps are then upsampled accordingly and overlaid on the original chest X-ray images. This supports in visualizing the areas that the proposed model deems important as given in Fig. 15. For interpretations, these visualizations are crucial since they allow clinical people to understand and verify the decision-making ability of the employed model. Thus, the highlighted areas as shown in Fig. 15 ensure that the model has focused correctly on relevant features. This results in improved trust and confidence in the automated diagnosis system.

#### Comparison with the recent existing works

The SEA-ResNet50 approach is finally compared with the state-of-the-art works, and it is listed in Table 4. This listed comparison summary includes the recently published and its associated research studies. And these summarized works have taken either CT or chest X-ray images as input. The comparison summary listed in Table 4 reveals that the proposed SEA-ResNet50 architecture together with the adaptive mish activation and ranger optimizer outperforms other recent approaches. This is due to the appropriate choice of transfer learning model with the right choice of activation and

optimization functions for the multiclass classification of the COVID-19 problem.

#### Discussion of the findings

The outline plot of the research aims in improving the performance of ResNet-50 architecture for the COVID-19 multiclass classification task as summarized next. Here, since the problem involves multiclass classification, the complexity and challenges involved are class imbalance, feature dimensionality, preprocessing methodologies, complex relationships between input features and target classes, inter-class variability, model selection and tuning, and evaluation metrics.

- The originality of the work lies in improving the ResNet50 model's performance for COVID-19 classification through the proposed SEA-ResNet50 deep learning architecture.
- This is done through the utilization and successful integration of the squeeze-and-excitation attention, adaptive mish activation and ranger optimization functions with the ResNet50 model.
- The study utilized the ResNet50 [46] as compared with the VGG16 [41], Xception [42], ResNet18 [45], and DenseNet121 [51].
- The ResNet50 model is chosen in this research because it provides a balance between depth and computational complexity. In addition, the model is deeper than VGG16 and ResNet18 but still compu-

**Table 4** Comparative summary of the proposed approach with the recently published studies for COVID-19 classification problem

Research work	Dataset	Architecture	Class	Classification accuracy (%)
[21]—Asif Iqbal et al	Chest X-ray images taken from GitHub and Kaggle	CoroNet	4 3	89.6 95.0
[59]—Tanvir et al	Chest X-ray images collected from Guangzhou Medical Center, China, and Sylhet Medical College, Bangladesh	CovXNet	2 4 3 2	99.0 90.3 89.6 94.7
[60]—Tulin et al	Chest X-ray images of CohenJP and Wang databases	DarkCovidNet	3 2	98.0 87.0
[61]—Marko et al	ChestXRay2017 dataset	ResNetCOVID19	3	94.1
[62]- Prabira et al	Chest X-ray images of CohenJP database	ResNet50+	2	95.3
[63]—Shahin et al	COVID-CT-MD dataset	COVID-Fact	3	90.8
[64]—Xiaowei et al	Own dataset collected from three distinct hospitals	ResNet with Attention	3	86.7
[65]—Himadri et al	Chest X-ray images taken from GitHub and Kaggle	Deep Neural Architecture	2	96.13
[66]—Emin Sahin	Chest X-ray images taken from Kaggle	Deep Convolutional Architecture	2	96.7
[67]—Umair et al	COVID X-ray images taken from Kaggle	CODISC-CNN	2 3 4	97.0 89.0 84.0
[68] – Patro et al	Chest X-ray images taken from GitHub and Kaggle	SConvNet50	2	97.33
Proposed Architecture	Chest X-ray images taken from GitHub and Kaggle	SEA-ResNet50 Model with Adaptive Mish Activation and Ranger Optimizer	2 4	99.29 98.38

tationally efficient as compared with more complex architectures such as DenseNet121 and Xception models. This balance makes the ResNet50 model a good choice for the employed COVID-19 classification problem.

- The above CAD model is evaluated in this study for both binary (Normal Vs COVID-19) and multiclass (Normal, Pneumonia, Lung\_Opacity, and COVID-19) classification tasks.

- As compared with recent research works, the proposed CAD approach outperforms and thereby establishes the novelty of the framework.

### Limitations of the study

Various real-time social crises always expect more promising outcomes through imaging analysis and algorithms using artificial intelligence. For all problems and the sake of saving lives, researchers all around the world are developing models for several real-world tasks. In this manner, the proposed CAD model is implemented successfully, and precise outcomes have been attained. However, as from Table 2, all the employed and proposed deep learning architectures have struggled to discriminate the chest X-ray images belonging to Pneumonia (Class\_2) during multiclass classification. This implies that the proposed DL architecture further requires to be improved to attain more promising results i.e., minimizing the risk of discriminating pneumonia inputs as either of the other three classes. Subsequently, the employed dataset

contains lesser image inputs of pneumonia as compared with other classes. Accordingly, this problem could be tackled in the future extension of this work.

### Conclusion and future work

In today's era, the risk to human lives is increasing due to several diseases. However, the research community has developed different promising algorithms for saving human lives. The negative impact on human lives was severe due to coronavirus and its subsequent pandemic. The research in this paper intended to design a CAD for both multiclass and binary classification of COVID-19 severities. The classes considered in our problem are Normal, Pneumonia, Lung\_Opacity, and COVID. For this, a combined dataset of chest X-ray image data are taken for assessment. The proposed method starts with the robust and popular transfer learning architecture, ResNet50. The model's classification performance is further improved using the Squeeze-and Excitation Attention (SEA) modules, adaptive Mish activation function, and Ranger optimizer. Different experiments with appropriate ablation studies were conducted for ensuring the robustness of the proposed SEA-ResNet50 model. A new learnable and adaptive parameter ( $\alpha$ ) is introduced in the activation function for attaining robust classification. In addition, the experimentation using VGG16, Xception, ResNet18, ResNet50, and DenseNet121 models is done for the comparative analysis of the proposed CAD model. The results of the research are validated using Kappa validation. Accordingly, the attained outcomes reveal that

the proposed approach SEA-ResNet50 outperformed the performance of other DL models in both binary and multiclass classification of COVID-19 severity. That is the proposed model provided maximum overall classification accuracies of 98.38% (multiclass) and 99.29% (binary) as compared with the existing works. The robustness of the proposed model is evaluated with respective kappa validation and the scores obtained are 0.975 and 0.98 for multiclass and binary classification of COVID-19 severity. The direction of our future work will be implementing the proposed SEA-ResNet50 model with different chest X-ray image datasets with distinct image preprocessing approaches. Also, future research will include the evaluation of the proposed model with both X-ray and CT images for COVID-19 diagnosis.

#### Acknowledgements

Not applicable.

#### Authors' contributions

S.C.S.R and B.N took care of the review of literature and methodology. V.C and V.K.V have done the formal analysis, data collection and investigation. M.T.R has done the initial drafting and statistical analysis. S.G has supervised the overall project. All the authors of the article have read and approved the final article.

#### Funding

This research received no external funding.

#### Availability of data and materials

The data that support the findings of this study are included in the manuscript.

#### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

NA.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Department of Electronics and Communication Engineering, Bannari Amman Institute of Technology, Sathyamangalam, India. <sup>2</sup>School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Vellore 632014, India. <sup>3</sup>Department of Computer Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, India. <sup>4</sup>Department of Computer Science and Engineering, JAIN (Deemed-to-Be University), Bengaluru 562112, India. <sup>5</sup>Adama Science and Technology University, Adama 302120, Ethiopia.

Received: 4 April 2024 Accepted: 6 August 2024

Published online: 09 August 2024

#### References

- Kumar A, Singh R, Kaur J, Pandey S, Sharma V, Thakur L, Sati S, Mani S, Asthana S, Sharma TK, Chaudhuri S. Wuhan to world: the COVID-19 pandemic. *Front Cell Infect Microbiol*. 2021;11:596201.
- Alimohamadi Y, Sepandi M, Taghdir M, Hosamirudisari H. Determine the most common clinical symptoms in COVID-19 patients: a systematic review and meta-analysis. *J Prev Med Hyg*. 2020;61(3):E304.
- Chaimayo C, Kaewnaphan B, Tanlieng N, Athipanyasilp N, Sirijatuphat R, Chayakulkeeree M, Angkasekwinai N, Sutthent R, Puangpunngam N, Tharmviboonsri T, Pongraweevan O. Rapid SARS-CoV-2 antigen detection assay in comparison with real-time RT-PCR assay for laboratory diagnosis of COVID-19 in Thailand. *Virology journal*. 2020;17:1–7.
- Schoy A, Anantharajah A, Bodéus M, Kabamba-Mukadi B, Verroken A, Rodriguez-Villalobos H. Low performance of rapid antigen detection test as frontline testing for COVID-19 diagnosis. *J Clin Virol*. 2020;129:104455.
- Borakati A, Perera A, Johnson J, Sood T. Diagnostic accuracy of X-ray versus CT in COVID-19: a propensity-matched database study. *BMJ Open*. 2020;10(11):e042946.
- Shi W, Tong L, Zhu Y, Wang MD. COVID-19 automatic diagnosis with radiographic imaging: Explainable attention transfer deep neural networks. *IEEE J Biomed Health Inform*. 2021;25(7):2376–87.
- Nayak SR, Nayak DR, Sinha U, Arora V, Pachori RB. Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomed Signal Process Control*. 2021;64:102365.
- Soares LP, Soares CP. Automatic detection of covid-19 cases on x-ray images using convolutional neural networks. 2020. arXiv preprint [arXiv:2007.05494](https://arxiv.org/abs/2007.05494).
- Das AK, Ghosh S, Thunder S, Dutta R, Agarwal S, Chakrabarti A. Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. *Pattern Anal Appl*. 2021;24:1111–24.
- Monshi MMA, Poon J, Chung V, Monshi FM. CovidXrayNet: Optimizing data augmentation and CNN hyperparameters for improved COVID-19 detection from CXR. *Comput Biol Med*. 2021;133:104375.
- Rajpal S, Lakhyani N, Singh AK, Kohli R, Kumar N. Using handpicked features in conjunction with ResNet-50 for improved detection of COVID-19 from chest X-ray images. *Chaos, Solitons Fractals*. 2021;145:110749.
- Kumar A, Tripathi AR, Satapathy SC, Zhang YD. SARS-Net: COVID-19 detection from chest x-rays by combining graph convolutional network and convolutional neural network. *Pattern Recogn*. 2022;122:108255.
- Alimadadi A, Aryal S, Manandhar I, Munroe PB, Joe B, Cheng X. Artificial intelligence and machine learning to fight COVID-19. *Physiol Genomics*. 2020;52(4):200–2.
- Saba L, Agarwal M, Patrick A, Puvvula A, Gupta SK, Carriero A, Laird JR, Kitas GD, Johri AM, Balestrieri A, Falaschi Z. Six artificial intelligence paradigms for tissue characterisation and classification of non-COVID-19 pneumonia against COVID-19 pneumonia in computed tomography lungs. *Int J Comput Assist Radiol Surg*. 2021;16:423–34.
- Wang L, Lin ZQ, Wong A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. *Sci Rep*. 2020;10(1):19549.
- Farooq M, Hafeez A. Covid-resnet: a deep learning framework for screening of covid19 from radiographs. 2020. arXiv preprint [arXiv:2003.14395](https://arxiv.org/abs/2003.14395).
- Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Med Image Anal*. 2020;65:101794.
- Ucar F, Korkmaz D. COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med Hypotheses*. 2020;140:109761.
- Novitasari DCR, Hendradi R, Caraka RE, Rachmawati Y, Fanani NZ, Syarifudin A, Toharudin T, Chen RC. Detection of COVID-19 chest X-ray using support vector machine and convolutional neural network. *Commun Math Biol Neurosci*. 2020;2020:Article-ID.
- Rajaraman S, Antani S. Weakly labeled data augmentation for deep learning: a study on COVID-19 detection in chest X-rays. *Diagnostics*. 2020;10(6):358.
- Khan AI, Shah JL, Bhat MM. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput Methods Programs Biomed*. 2020;196:105581.
- Pereira RM, Bertolini D, Teixeira LO, Silla CN Jr, Costa YM. COVID-19 identification in chest X-ray images on flat and hierarchical classification scenarios. *Comput Methods Programs Biomed*. 2020;194:105532.
- Hussain L, Nguyen T, Li H, Abbasi AA, Lone KJ, Zhao Z, Zaib M, Chen A, Duong TQ. Machine-learning classification of texture features of portable

- chest X-ray accurately classifies COVID-19 lung infection. *Biomed Eng Online*. 2020;19:1–18.
24. Ukwuoma CC, Qin Z, Agbesei VK, Cobbinah BM, Yussif SB, Abubakar HS, Lemessa BD. Dual\_Pachi: Attention-based dual path framework with intermediate second order-pooling for Covid-19 detection from chest X-ray images. *Comput Biol Med*. 2022;151:106324.
  25. Khero K, Usman M, Fong A. Deep learning framework for early detection of COVID-19 using X-ray images. *Multimedia Tools and Applications*. 2024;83(3):6883–908.
  26. Serte S, Demirel H. Deep learning for diagnosis of COVID-19 using 3D CT scans. *Comput Biol Med*. 2021;132:104306.
  27. Bahgat WM, Balaha HM, AbdulAzeem Y, Badawy MM. An optimized transfer learning-based approach for automatic diagnosis of COVID-19 from chest x-ray images. *PeerJ Computer Science*. 2021;7:e555.
  28. Song Y, Zheng S, Li L, Zhang X, Zhang X, Huang Z, ..., Yang Y. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Trans Comput Biol Bioinform*. 2021;18(6):2775–2780.
  29. Azeem M, Javaid S, Khalil RA, Fahim H, Althobaiti T, Alsharif N, Saeed N. Neural Networks for the Detection of COVID-19 and Other Diseases: Prospects and Challenges. *Bioengineering*. 2023;10(7):850.
  30. Mercaldo F, Belfiore MP, Reginelli A, Brunese L, Santone A. Coronavirus COVID-19 detection by means of explainable deep learning. *Sci Rep*. 2023;13(1):462.
  31. Tembhumne J. Classification of COVID-19 patients from HRCT score prediction in CT images using transfer learning approach. *J Electric Syst Info Technol*. 2024;11(1):1–13.
  32. Gopatoti A, Vijayalakshmi P. MTMC-AUR2CNet: Multi-textural multi-class attention recurrent residual convolutional neural network for COVID-19 classification using chest X-ray images. *Biomed Signal Process Control*. 2023;85:104857.
  33. Gopatoti A, Vijayalakshmi P. CXGNet: A tri-phase chest X-ray image classification for COVID-19 diagnosis using deep CNN with enhanced grey-wolf optimizer. *Biomed Signal Process Control*. 2022;77:103860.
  34. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Bin Mahbub Z, Islam KR, Khan MS, Iqbal A, Al-Emadi N, Reaz MBI, Islam TI. Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access*. 2020. <http://arxiv.org/abs/2003.13145>. Accessed 05 Dec 2023.
  35. Maguolo G, Nanni L. A critic evaluation of methods for COVID-19 automatic detection from X-Ray images. 2020. <http://arxiv.org/abs/2004.12823>. Accessed 05 Dec 2023.
  36. Whybra P, Zwanenburg A, Andrearczyk V, Schaer R, Apte AP, Ayotte A, Baheti B, Bakas S, Bettinelli A, Boellaard R, Boldrini L. The image biomarker standardization initiative: Standardized convolutional filters for reproducible radiomics and enhanced clinical insights. *Radiology*. 2024;310(2):e231319.
  37. Sannasi Chakravarthy SR, Rajaguru H. A novel improved crow-search algorithm to classify the severity in digital mammograms. *Int J Imaging Syst Technol*. 2021;31(2):921–54.
  38. Chakravarthy SS, Rajaguru H. Automatic detection and classification of mammograms using improved extreme learning machine with deep learning. *Irbm*. 2022;43(1):49–61.
  39. Thepade SD, Pardhi PM. Contrast enhancement with brightness preservation of low light images using a blending of CLAHE and BPDHE histogram equalization methods. *Int J Inf Technol*. 2022;14(6):3047–56.
  40. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition. 2009. <https://doi.org/10.1109/cvprw.2009.5206848>.
  41. Sannasi Chakravarthy SR, Bharanidharan N, Rajaguru H. Multi-deep CNN based experimentations for early diagnosis of breast cancer. *IETE J Res*. 2023;69(10):7326–41.
  42. Chollet F. Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2017. p. 1251–8.
  43. Wu D, Wang Y, Xia ST, Bailey J, Ma X. Skip connections matter: on the transferability of adversarial examples generated with resnets. 2020. arXiv preprint [arXiv:2002.05990](https://arxiv.org/abs/2002.05990).
  44. Oyedotun OK, Al Ismaeil K, Aouada D. Why is everyone training very deep neural network with skip connections? In: IEEE Transactions on Neural Networks and Learning Systems. 2022.
  45. Chakravarthy SS, Bharanidharan N, Rajaguru H. Deep Learning-Based Metaheuristic Weighted K-Nearest Neighbor Algorithm for the Severity Classification of Breast Cancer. *Irbm*. 2023;44(3):100749.
  46. Mukti IZ, Biswas D. Transfer Learning Based Plant Diseases Detection Using ResNet50. 2019 4th International Conference on Electrical Information and Communication Technology (EICT). 2019. <https://doi.org/10.1109/eict48899.2019.9068805>.
  47. Talaat FM, El-Sappagh S, Alnowaiser K, Hassan E. Improved prostate cancer diagnosis using a modified ResNet50-based deep learning architecture. *BMC Med Inform Decis Mak*. 2024;24(1):23.
  48. Li MH, Yu Y, Wei H, Chan TO. Classification of the qilou (arcade building) using a robust image processing framework based on the Faster R-CNN with ResNet50. *J Asian Architect Build Engineer*. 2024;23(2):595–612.
  49. Chen Y, Liu J, Jiang P, Jin Y. A novel multilevel iterative training strategy for the ResNet50 based mitotic cell classifier. *Comput Biol Chem*. 2024;110:108092.
  50. Chen Y, Wang L, Ding B, Shi J, Wen T, Huang J, Ye Y. Automated Alzheimer's disease classification using deep learning models with Soft-NMS and improved ResNet50 integration. *J Rad Res Appl Sci*. 2024;17(1):100782.
  51. Huang Z, Zhu X, Ding M, Zhang X. Medical image classification using a light-weighted hybrid neural network based on PCANet and DenseNet. *IEEE Access*. 2020;8:24697–712.
  52. Hu J, Shen L, Sun G. Squeeze-and-excitation networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2018. p. 7132–41.
  53. Misra D. Mish: A self regularized non-monotonic activation function. 2019. arXiv preprint [arXiv:1908.08681](https://arxiv.org/abs/1908.08681).
  54. Wright L, Demeure N. Ranger21: a synergistic deep learning optimizer. 2021. arXiv preprint [arXiv:2106.13731](https://arxiv.org/abs/2106.13731).
  55. Sannasi Chakravarthy SR, Rajaguru H. Detection and classification of microcalcification from digital mammograms with firefly algorithm, extreme learning machine and non-linear regression models: A comparison. *Int J Imaging Syst Technol*. 2020;30(1):126–46.
  56. Sannasi Chakravarthy SR, Rajaguru H. Performance analysis of ensemble classifiers and a two-level classifier in the classification of severity in digital mammograms. *Soft Comput*. 2022;26(22):12741–60.
  57. Arrieta AB, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, García S, Gil-López S, Molina D, Benjamins R, Chatila R. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information fusion*. 2020;58:82–115.
  58. Kim JK, Jung S, Park J, Han SW. Arrhythmia detection model using modified DenseNet for comprehensible Grad-CAM visualization. *Biomed Signal Process Control*. 2022;73:103408.
  59. Mahmud T, Rahman MA, Fattah SA. 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*. 2020;122:103869.
  60. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med*. 2020;121:103792.
  61. Arsenovic M, Sladojevic S, Orcic S, Anderla A, Sladojevic M. Detection of COVID-19 cases by utilizing deep learning algorithms on X-ray images. 2020.
  62. Sethy PK, Behera SK. Detection of coronavirus disease (covid-19) based on deep features. 2020.
  63. Heidarian S, Afshar P, Enshaei N, Naderkhani F, Rafiee MJ, Babaki Fard F, Samimi K, Atashzar SF, Oikonomou A, Plataniotis KN, Mohammadi A. Covid-fact: A fully-automated capsule network-based framework for identification of covid-19 cases from chest ct scans. *Frontiers in Artificial Intelligence*. 2021;4:598932.
  64. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J, Lang G. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering*. 2020;6(10):1122–9.
  65. Mukherjee H, Ghosh S, Dhar A, Obaidullah SM, Santosh KC, Roy K. Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays. *Appl Intell*. 2021;51:2777–89.
  66. Sahin ME. Deep learning-based approach for detecting COVID-19 in chest X-rays. *Biomed Signal Process Control*. 2022;78:103977.

67. Hafeez U, Umer M, Hameed A, Mustafa H, Sohaib A, Nappi M, Madni HA. A CNN based coronavirus disease prediction system for chest X-rays. *J Ambient Intell Humaniz Comput.* 2023;14(10):13179–93.
68. Patro KK, Allam JP, Hammad M, Tadeusiewicz R, Pławiak P. SCovNet: A skip connection-based feature union deep learning technique with statistical approach analysis for the detection of COVID-19. *Biocybernet Biomed Engineer.* 2023;43(1):352–68.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.