

## ORIGINAL RESEARCH

# An effective screening of COVID-19 pneumonia by employing chest X-ray segmentation and attention-based ensemble classification

Abu Sayeed<sup>1</sup> | Nasif Osman Khansur<sup>1</sup> | Azmain Yakin Srizon<sup>1</sup> |  
Md. Farukuzzaman Faruk<sup>1</sup> | Salem A. Alyami<sup>2</sup> | AKM Azad<sup>2</sup> | Mohammad Ali Moni<sup>3</sup> 

<sup>1</sup>Department of Computer Science & Engineering, Faculty of Electrical & Computer Engineering, Rajshahi University of Engineering and Technology, Kazla, Rajshahi, Bangladesh

<sup>2</sup>Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arabia

<sup>3</sup>AI & Digital Health Technology, Artificial Intelligence and Cyber Futures Institute, Charles Sturt University, Bathurst, NSW, Australia

**Correspondence**

Mohammad Ali Moni, AI & Digital Health Technology, Artificial Intelligence and Cyber Futures Institute, Charles Sturt University, Bathurst, NSW 2795, Australia.

Email: [mmoni@csu.edu.au](mailto:mmoni@csu.edu.au); [m.moni@uq.edu.au](mailto:m.moni@uq.edu.au)

Salem A. Alyami, Department of Mathematics and Statistics, Faculty of Science, Imam Mohammad Ibn Saud Islamic University (IMSIU), Riyadh, Saudi Arab.

Email: [saalyami@imamu.edu.sa](mailto:saalyami@imamu.edu.sa)

**Funding information**

Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU), Grant/Award Number: RPP2023003

**Abstract**

Quick and accurate diagnosis of COVID-19 is crucial in preventing its transmission. Chest X-ray (CXR) imaging is often used for diagnosis, however, even experienced radiologists may misinterpret the results, necessitating computer-aided diagnosis. Deep learning has yielded favourable results previously, but overfitting, excessive variance, and generalization errors may occur due to noise and limited datasets. Ensemble learning can improve predictions by using robust techniques. Therefore, this study, proposes two-fold strategy that combines advanced and robust algorithms, including DenseNet201, EfficientNetB7, and Xception, to achieve faster and more accurate COVID-19 detection. Segmented lung images were generated from CXR images using the residual U-Net model, and two attention-based ensemble neural networks were used for classification. The COVID-19 radiography dataset was used to evaluate the proposed approach, which achieved an accuracy of 98.21%, 93.4%, and 89.06% for two, three, and four classes respectively which outperformed previous studies by a significant margin considering COVID, viral pneumonia, and lung opacity simultaneously. Despite the similarity in CXR images of COVID, pneumonia, and lung opacity, the proposed approach achieved 89.06% accuracy, demonstrating its ability to recognize distinguishable features. The developed algorithm is expected to have applications in clinics for diagnosing different diseases using X-ray images.

## 1 | INTRODUCTION

A cluster of pneumonia cases reported in Wuhan, China, has swiftly spread well over the world within a short period of time, according to a study from late 2019 [1–3]. Since its inception, a newly discovered coronavirus has caused havoc and eventually become a pandemic [4]. The sickness, which is termed COVID-19 according to the World Health Organization (WHO), is caused by the SARS-CoV-2 virus. Most often, those who contract the virus experience a mild-to-moderate respiratory infection and recover without any complications, but older

people requiring intensive medical support for pre-conditions including cancer, diabetes, cardiovascular disease, or chronic respiratory diseases are more at risk [3]. Most coronavirus patients will experience a common cold and fever, according to studies, while only a small number will show no symptoms.

The best resilience against the infection has been demonstrated within adults, but as a drawback, they are more apt to transfer it than adolescents [5].

More than 585 million confirmed cases with 6.4 million fatalities and 560 million recoveries had been reported as of August 2022, in roughly 228 nations [6]. Coronavirus has also

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Authors. *IET Image Processing* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

impacted the medical staff who are on the front lines of the battle against the coronavirus epidemic. According to a study, this illness affects the emotional, physical, and social health of medical workers [7]. These days, antigen testing and reverse transcription polymerase chain reaction (RT-PCR) testing, which can detect the active SARS-CoV-2 protein as well as RNA from respiratory specimens, most regularly through oro- or nasopharyngeal swabs, are the most renowned clinical screening methods for the detection of COVID-19 active cases [7].

Although these strategies are the gold standard in detection, there are debatable key criteria. Antigen testing seems to be less expensive and causes a tremor, but occasionally produces erroneous results because of its low sensitivity to the ratios of symptomatic and asymptomatic patients [8]. The antigen test is sensitive when the infection rate is at its peak, i.e. the highest concentration of SARS-CoV-2 proteins. Since there may be insufficient cellular material or inadequate extractions in the sample that was obtained from the infected person, which may be resulted in ambiguous diagnostics. Moreover, RT-PCR testing takes a long time to assess the results and is also much more difficult and laborious. Although the initial barriers to the provision of kits for the aforementioned approaches have already been alleviated, the drawback in terms of sensitivity may still exist for samples that are collected less than optimally or in patients with a low viral load [9]. More specifically, 93% of samples from broncho-alveolar lavage, 63% of samples from sputum, and only 32% of samples from throat swabs are sensitive enough for RT-PCR testing [10].

CXR and CT imaging examinations are two significant alternate standard screening techniques that are utilized [11]. These techniques are used by radiologists as visual indicators for diagnosing viral infections. According to certain experimental research on the correlation between patients with confirmed COVID-19 pneumonia clinical circumstances and chest CT findings, these patients have typical imaging characteristics that make them amenable to early disease scope screenings [11]. The major challenge faced by CT imaging is that after monitoring symptoms for a prolonged period, the initial CT report appears normal [12]. According to one study, chest CT has an excellent sensitivity of up to 90% but a comparatively lower specificity of between 25% in 83% of adult populations, while the accuracy in paediatric populations appears to be quite poor due to insufficient evidence [9]. In these circumstances, CXR imaging meets the fundamental requirements for better screening of the COVID-19 population worldwide, particularly in resource-constrained and severely afflicted areas [9]. The rapid triage of patients with suspected COVID-19 is made possible by the availability of CXR imaging equipment, which is easily available to many healthcare centers and may be used inside an isolation room, thus approving its portability [7, 10].

CXR imaging has manageable maintenance costs. Radiologists with advanced training and expertise are required for COVID-19 detection if the CXR image analysis is used. Yet, compared to the whole population screened, there are, sadly, significantly fewer radiologists or specialized doctors. Based on data from 29 nations, a recent study found that there were

3.52 active doctors for every thousand inhabitants on average. The greatest number was 5.24 doctors per 1,000 people in Austria, and the lowest value was 2.18 doctors per 1,000 people in Colombia [13]. Due to its capacity to function in the face of massive datasets that outperform humanistic potential, AI-based techniques using deep learning have proven to be successful and have had a positive impact on the medical field. In relation to COVID-19 identification utilizing deep convolution networks, numerous academics from around the world are experimenting with different datasets and releasing articles. Although these methods have shown promise in medium-to-light datasets, they have rarely been approved as solutions that are suitable for production. Medical imaging does not currently have enough data to train deep learning models. New deep learning models can't generalize to clinical data that hasn't been seen, even if the training data is small. Additionally, medical images need to be manually labelled and annotated, which takes time and money and is expensive [14]. Transfer learning may offer great alternatives for locating precisely labelled data, as the models' capacity for handling large numbers of data is severely constrained, and it also eliminates the necessity for human annotators [14, 15]. A medical analysis is nothing unusual, but transfer learning has recently proved quite successful in a variety of visual tasks. When it comes to automatic COVID-19 identification, multiple state-of-the-art pre-trained models that apply transfer learning outperform novel convolutional structures that are created from scratch [10, 16, 17].

In our study, we have presented a deep convolutional neural network architecture using segmented chest X-ray images for automatic COVID-19 recognition in a computer-aided diagnosis model. In order to accomplish improved prediction accuracy, we used an ensemble of three pre-trained models (DenseNet201, Xception, and EfficientNetB7) after using the ResUNet model to segment the lung section from X-ray images. To the best of our knowledge, the model outperformed many other state-of-the-art experimental studies, achieving a higher accuracy of 98.21% for two classes, 93.4% for three classes, and 89.06% for four classes, respectively. In summary, our main contributions to this work are as follows:

- The model has an end-to-end hierarchical structure with fewer parameters, which requires fewer operations and saves memory space.
- Segmented lung region of raw CXR images have been used without performing any manual image augmentation process.
- Proposing an ensemble learning approach for achieving higher predictive accuracy and robustness.
- Large training dataset has been used with no manual feature extraction and preprocessing techniques.

The whole article is organized into different sections, i.e. a review of related works in Section 2, dataset description and model formulation in Section 3, evaluation of training, and test results with confusion matrices are discussed in Section 4. In Section 5, the future scope of this research is addressed and the conclusion is pointed out in Section 6.

## 2 | RELATED WORKS

The development of AI-based medical systems and the quick rise of COVID-19 have resulted in the widespread use of automatic computer-aided diagnosis. Numerous researchers from all around the globe are working on developing a quicker and more effective detection model. Binary or categorical classifications have been used in studies on the prediction of COVID-19 using CXR pictures; some of these studies handled raw data, while others used preprocessing techniques. In systematic investigations, the sizes of datasets are also crucial. [16] developed a model named DarkCovidNet based on DarkNet families that attained an accuracy of 87.02% for three classes. [10] tailored a deep neural network and defined that as COVID-NET for the early detection of COVID-19 on CXR images using a projection-expansion-projection-extension (PEPX) structure that has obtained an accuracy of 93.3%. In an experimental study, [18] performed scale-invariant feature transform (SIFT) & binary robust invariant scalable key-points (BRISK)-based deep learning models to predict COVID-19 infected patients, which acquired 96.6% accuracy. [19] proposed a complex model on MCFE-Net, which obtained an accuracy of 96.79% for three classifications, but the model faced a hardware issue in practical application. In a study, [20] achieved 90.3% accuracy with the augmentation and normalization of CXR images using ResNet50, ResNet152, and VGG16 models. [21] investigated how different enhancement strategies affected both plain and segmented images and presented a modified U-net model for lung segmentation, obtaining accuracy and intersection over union (IoU) of 98.63% and 94.3%, respectively. They investigated six pre-trained models (i.e. ResNet18, ResNet50, ResNet101, InceptionV3, DenseNet201, and ChexNet - some of their details can be found in the Section 5) and found that DenseNet201 outperformed the rest with an accuracy of 95.11% for segmenting gamma-corrected lung images. Another investigation revealed that accuracy rates for classifications into three and four classes using the CoroDet model were 94.2% and 91.2%, respectively [22]. Using an ensemble of CNN and ResNet architecture, [23] performed the classification on three classes and achieved 93.1% and 90% accuracy, respectively. For the prior architectures for four classes, they also attained an accuracy of 77.2% and 81.5%. [24] used explainable artificial intelligence approaches to quantify the impact of segmentation, scoring 98.2% on the dice and then 83% on the F1-score for COVID identification with multiclass classification. [25] presented the Xception-based architecture, namely CoroNet to carry out the multiclass classification with an overall accuracy of 89.6%. DenseNet201 and MobileNet are used to extract features, and then selected features are submitted to LightGBM for multiclass classification, yielding an accuracy measure of 91.11%, according to a method described by [26]. In proposing a model called CovXnet, [27] noted an accuracy of 92.02%. In the COVID-19FclNet9 model, [28] reported an accuracy of 89.96%.

In an experiment, [29] used a model that included the weights of VGG16 and InceptionV3, which performed well

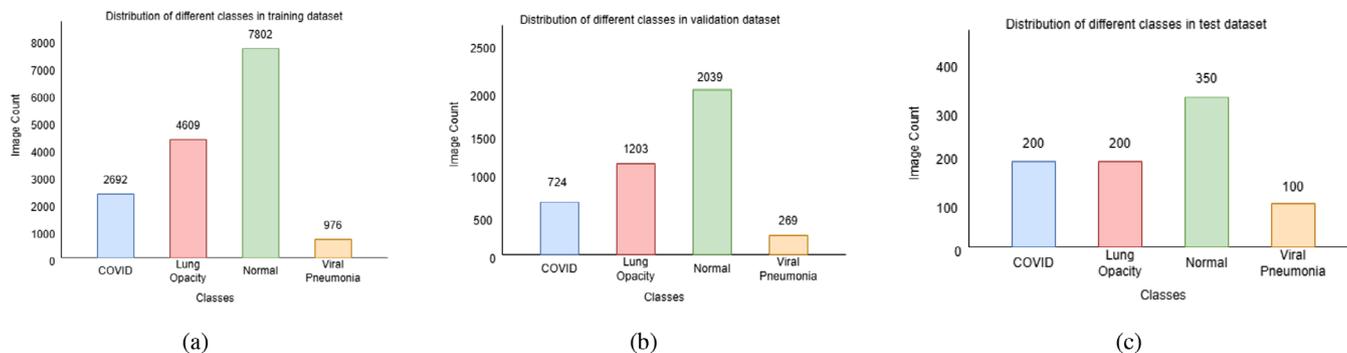
for binary and three-class classification but poorly for four-class classification, achieving just 62.5% accuracy. By applying the augmentation with VGG16 for the automatic classification of three different classes, [30] obtained an accuracy of 83.6%. Following the application of a deep learning framework based on VGG19, [31] suggested a semi-automated image preprocessing model to build a reliable dataset by removing noise and undesired features. The experiment's precision was 83%, while the experiment's sensitivity and f1-score were both 80%. [32] applied a VGG16-based model and found 84.10% accuracy for multiclass classification. [33] used CT images to screen for coronaviruses while taking into account ResNet and a location awareness system, resulting in an accuracy of 86.7%. Another experiment was carried out on CT scans by [34], who proposed the CovNet model and attained an accuracy of 90% for three classes. A concatenation of Xception and ResNet50V2 that was introduced by [35] yielded an overall accuracy of 91.4%, better than the performance of the separate models. [36] used MobileNetV2 and VGG19 to achieve accuracy of 98% and 95% in binary classifications, respectively. [37] reported a 98.7% accuracy rate for two classes using CNN-based transfer learning and a generative adversarial network. [38] discovered 87.86% accuracy for multiclass classification and 98.97% accuracy for binary classes in a different investigation.

## 3 | DATASET DESCRIPTION

Since COVID-19 is a novel disease, the relevant datasets are still evolving. The majority of publicly accessible datasets were created by compiling CXR images from several public image repositories. This article introduces the materials and dataset utilized in the study. The training, validation, and testing dataset distribution among the different classes are also discussed as illustrated in Figure 1.

In our study, we used a dataset publicly available on Kaggle named 'COVID-19 Radiography Database', the winner of the COVID-19 dataset award by the Kaggle Community [21, 39, 40]. The dataset contains CXR images for COVID-19-positive cases, normal, lung opacity, and viral pneumonia images along with their corresponding lung masks. A comprehensive amount of information, including the patient's demographics, the type of projection used, and the acquisition parameters for the imaging study, are provided together with the images that were annotated and validated by a team of expert radiologists. [19]

The database consists of 3,616 COVID-19-positive cases, with 10,192 normal, 6,012 lung opacity, and 1,345 viral pneumonia chest X-ray images along with their corresponding ground truth masks. All of the images have a resolution of 299×299 pixels and are stored as portable network graphics (PNG) files. While performing the segmentation technique, we chose 15,873 images belonging to 4 classes for training and 5,292 images for validation including the same number of segmentation masks. The segmented images are then further divided into three different categories. Figure 1 shows the distribution of different classes in the training, validation, and test datasets. Table 1



**FIGURE 1** Distribution of different classes in (a) training, (b) validation, and (c) test dataset.

**TABLE 1** Dataset counts for training, validation, and testing.

	COVID	Lung opacity	Normal	Viral pneumonia
Train	2692	4609	7802	976
Validation	724	1203	2039	269
Test	200	200	350	100

shows the overall dataset count for training, validation, and test datasets.

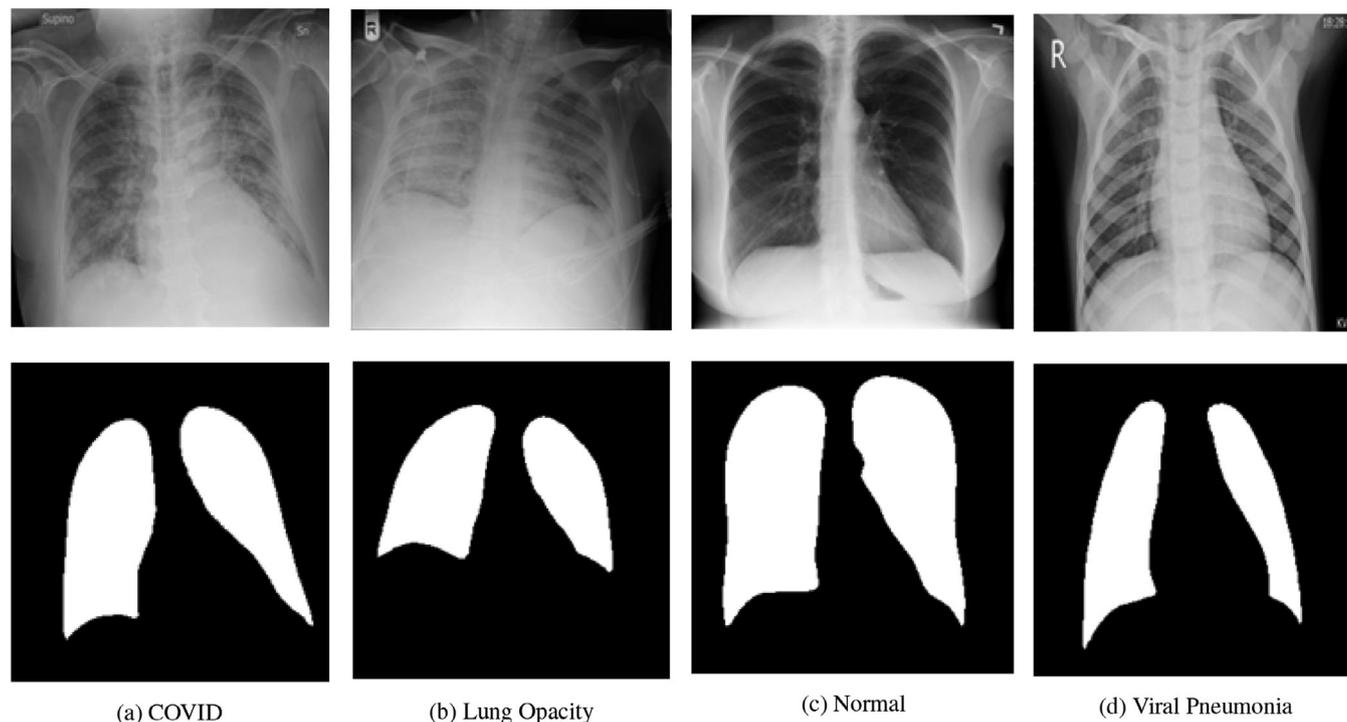
The amount of memory and computational resources needed to process larger images is directly impacted by the larger number of pixels that they contain. The overall processing time and memory needs can be greatly decreased by reducing the image size. This is particularly crucial when working with huge datasets. Additionally, shrinking the image size might direct the

model's attention to the areas of the image that are most informative [41]. We used raw images, so to enable the model to learn feature information more quickly, we reduced their size to a fixed resolution of 256×256 pixels for segmentation and 224×224 pixels for the classification study.

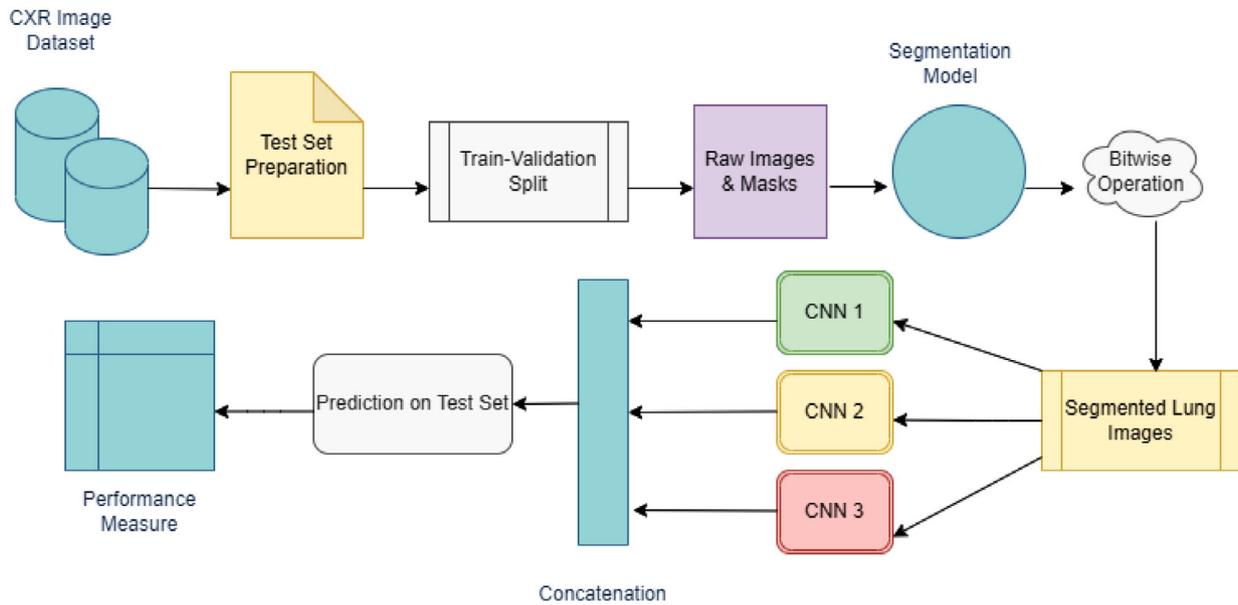
Figure 2 depicts the sample CXR images with their corresponding ground truth masks of normal, COVID-19, lung opacity, and viral pneumonia.

## 4 | METHODOLOGY

Convolutional neural networks have gained popularity as a result of their enhanced image classification capabilities. Convolutional layers and filters in a network, aid in the extraction of an image's spatial and temporal information where the fully connected layers make the model end-to-end trainable to classify



**FIGURE 2** Sample CXR images with their corresponding ground truth masks of (a) COVID, (b) lung opacity, (c) normal, and (d) viral pneumonia.



**FIGURE 3** The schematic diagram of our overall workflow.

the data into different classes. In a general sense, comparatively, for a small dataset, transfer learning may be helpful [46]. Transfer learning has recently been applied with effectiveness in a variety of practical applications, including manufacturing, medicine, and baggage screening [47–49]. This deep learning algorithm doesn't need to be created from scratch, which eliminates the need for a huge dataset and shortens the long learning period. Three different types of significant experiments have been used to carry out the entire process. The investigation required the creation of two distinct datasets. The training, and validation, plain CXR images for the first dataset, were derived straight from the original dataset, as opposed to the training, validation, and test segmented lung images for the second dataset, which were produced following the application of the segmentation approach. First, classification was carried out using 3 distinct pre-trained deep-learning models on plain CXR images. Individual models such as DenseNet201, Xception, and EfficientNetB7 were trained on our dataset, and the experimental results were recorded. Then, using a suitable segmentation model to pre-process our images by generating a specific lung region and a classification model to categorize the lung X-ray images, we set out to build a deep learning segmentation classification pipeline.

Our work has essentially been divided into two sections. The model was fed the raw CXR images and their related ground truth masks after dividing the dataset into training and validation. The predicted masks were then integrated with the images using a bit-wise AND operation. The segmented lung X-ray images were consequently produced. Three separate convolutional neural networks were then fed with the segmented images, which were then concatenated to make predictions on test data and evaluate the effectiveness of the suggested methodology. Figure 3 depicts the detailed workflow.

## 4.1 | Segmentation

In the literature, there are a number of U-net-based segmentation model variations. In the end, we discovered that the deep residual U-Net model, popularly known as ResUnet [50], produced the best results for us on the dataset. The model typically makes use of both residual deep learning and the U-Net architecture. The problem of vanishing gradients is alleviated by the residual connections in the residual U-Net. Additionally, the model has the capacity to learn more sophisticated and abstract representations of the input data. Compared to the conventional U-Net model, the residual U-Net has shown greater performance, with better accuracy, better retention of fine features, and improved segmentation boundaries. Therefore, we made the choice to continue using this segmentation module in our workflow. To perform the segmentation, images were resized to  $256 \times 256$ . This segmentation module, with a seven-level architecture, has three primary components theoretically. The encoder is considered the initial component of the architecture, which shrinks and compactifies the input images into smaller and more comprehensible representations. The decoder, or the final core part of this architecture, retrieves the representations into a pixel-by-pixel categorization, i.e. semantic segmentation. Between the encoder at the ResUNet's input and the decoder at the output, the classifier's second from the middle section acts as a link [50]. Figure 4 illustrates the model's complete architecture.

As an optimizer algorithm, the Adam optimizer was used, as it meets the requirements of momentum and root mean squared propagation (RMSprop), delivers faster speed in terms of invariance of the magnitudes of parameter, and handles the gradient descent problem. In comparison, it costs less to really get better returns. To obtain more detailed information, which is frequently employed in segmentation tasks, the dice-coefficient

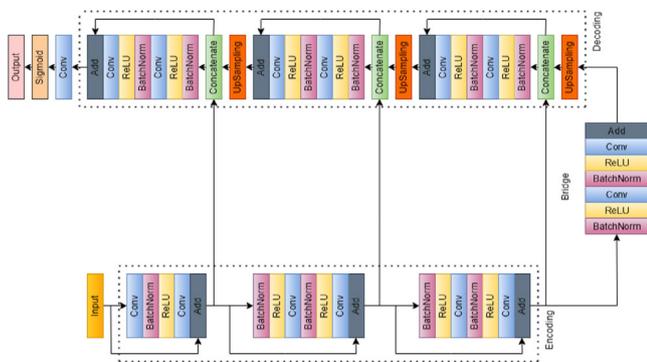


FIGURE 4 The architecture of ResUNet.

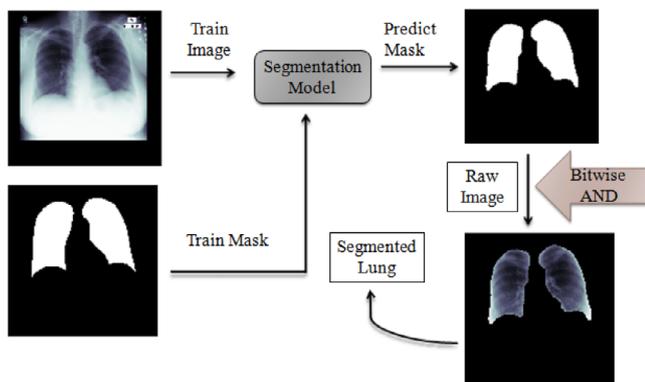


FIGURE 5 Complete visualization of the segmentation procedure.

loss function has been given priority over the simple cross-entropy. In the scenario of class imbalance, overall accuracy is not an appropriate statistic for semantic segmentation. Therefore, the metric ‘intersection over union’ (IoU) was preferred more. The complete visualization of the segmentation process is shown in Figure 5.

### 4.2 | Classification

Three deep learning models, DenseNet201 [51], EfficientNetB7 [52], and Xception [53], which were previously pre-trained on the ImageNet Database, have been trained on our processed dataset. These models have been applied to feature extraction because of their dense and scalable structures, which facilitate the effective learning of complicated patterns and rich

hierarchical features from images. Therefore, compared to computationally intensive networks trained from scratch, minimal computational resources and time are required which improves the performance of the classification. DenseNet201 is a convolutional neural network with 201 deep layers. It is composed of numerous dense blocks. Between layers adjacent to the input and those adjacent to the output, there are shorter connections. Transition layers between dense blocks are used to condense the spatial dimensions and manage the complexity of the model. These layers typically consist of an average pooling layer followed by a convolutional layer. The spatial data is combined into a single vector at the network’s end by a global average pooling layer. This model is more detailed, accurate, and easy to train as shown in Figure 6. An architecture of the model having three dense blocks has been shown.

The compound scaling strategy enhances the accuracy and effectiveness of the EfficientNetB7 model. It reduces the overall number of processes by scaling all width, depth, and resolution dimensions consistently. The broader layers are present between the skip connections, while the shortcut connections link the thin layers. On MBConv, squeeze-and-excitation optimization has been added. An architecture of the model having MBConv as the basic building block has been shown in Figure 7.

The improved or more extreme form of InceptionV3 is the Xception model with modified depth-wise separable convolution, which is essentially channel-wise  $n \times n$  spatial convolution. The network can learn identity mappings and substantially avoid the vanishing gradient issue because of the residual connections. It is significantly lighter because there are fewer connections. The architecture of the model is displayed in Figure 8.

For ensemble learning, we have presented two classification models. In the first model, after retrieving the inputs for the separate models’ convolutional parts, we applied a channel-wise attention mechanism over the extracted features and then directly concatenated them to feed the classification component. The overall method is depicted in Figure 9. This model was named proposed model 1. For the second model, we first individually trained the convolutional network for our dataset and saved the parameters. Then, using our own parameters rather than the ImageNet weights, we reloaded the models in accordance with the architecture illustrated in Figure 10. This model is named proposed model 2. Following that, we retrained each model using our dataset while freezing its upper layers and then concatenated the derived features. The features were sent to the fully connected or dense layers for classification once the attention mechanism was applied.

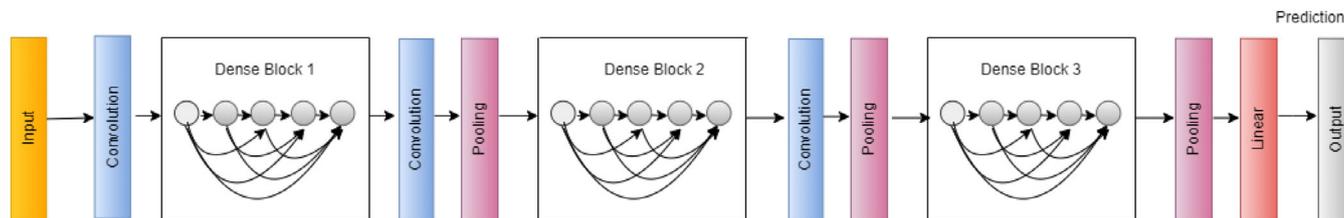


FIGURE 6 A three-dense-block deep DenseNet. The layers across two neighbouring blocks are known as transition layers, and they modify feature-map sizes through convolution and pooling.

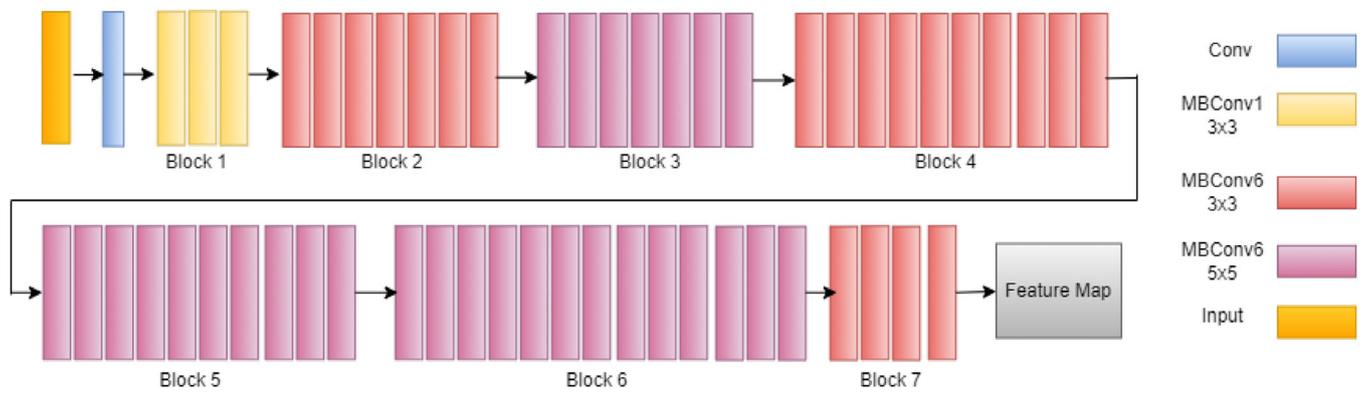


FIGURE 7 Visualization of EfficientNetB7's architecture using MBConv as a fundamental building block.

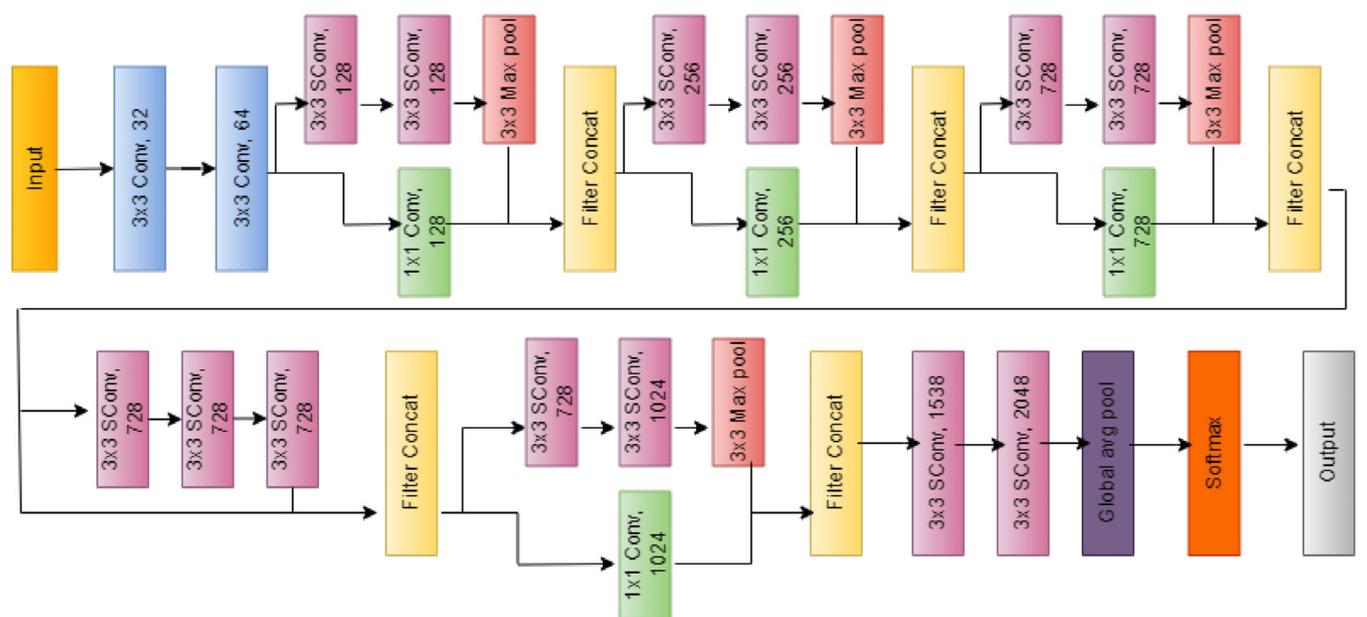


FIGURE 8 The Xception architecture.

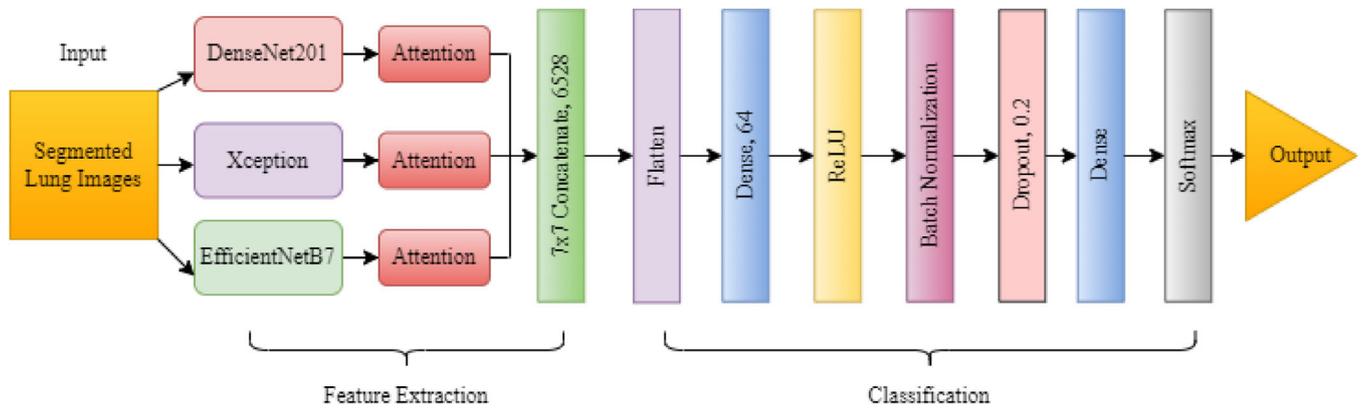


FIGURE 9 Proposed classification Model 1 using ensemble learning with attention mechanism.

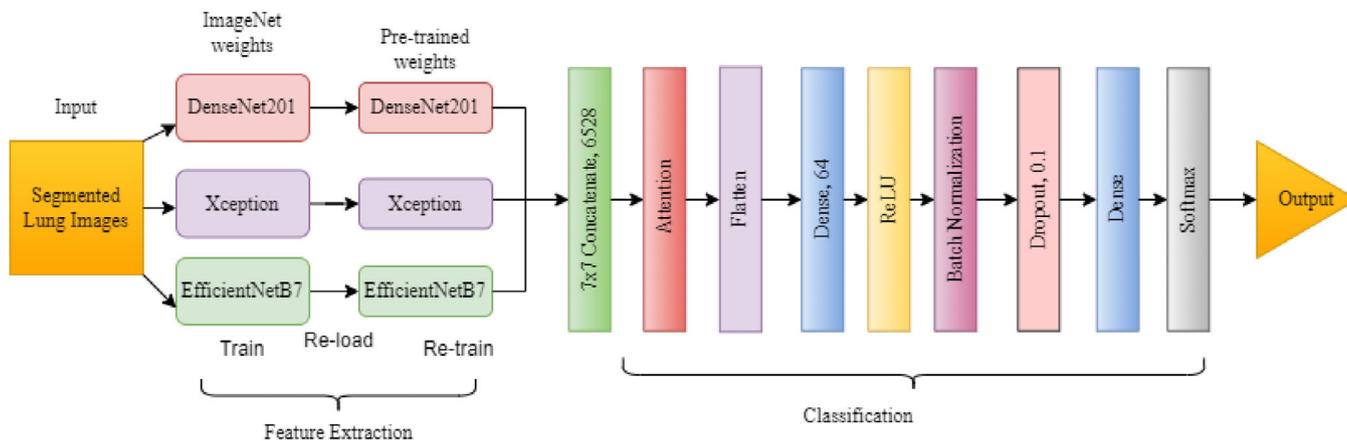


FIGURE 10 Proposed classification model 2 using ensemble learning with attention mechanism.

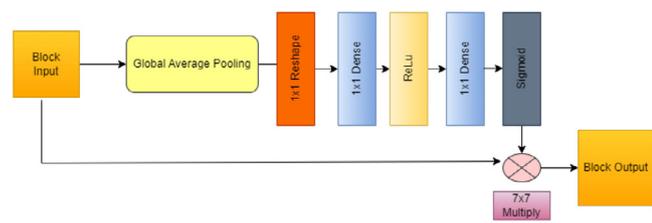


FIGURE 11 Block diagram of the SE attention module.

In Figure 11, a squeeze and excitation networks-based attention mechanism (SE-attention mechanism) is depicted, which is inspired by the unique attention mechanism published at CVPR 2018 [54]. Any basic architecture can be enhanced with this straightforward but effective add-on module to get improved performance with minimal computational overhead. The attention module primarily focuses on the channel inter-dependency and is utilized for feature re-calibration. It increases the network’s sensitivity to the object’s informational characteristics. An additional dense layer and a dropout layer were added to the end of the models in the process of being developed. These dense layers’ activation functions were all tuned to the ReLU activations. Each model had a dropout layer added at the end with a 10% dropout rate. This made it easier for each model to handle the limited size of our dataset and prevented overfitting. Our final layer included two neurons for categorization into two classes, three neurons for categorization into three classes, and four neurons for four classes, respectively, and each had a softmax activation function. Table 2 shows a summary of the proposed model 2’s architecture.

## 5 | RESULT & DISCUSSION

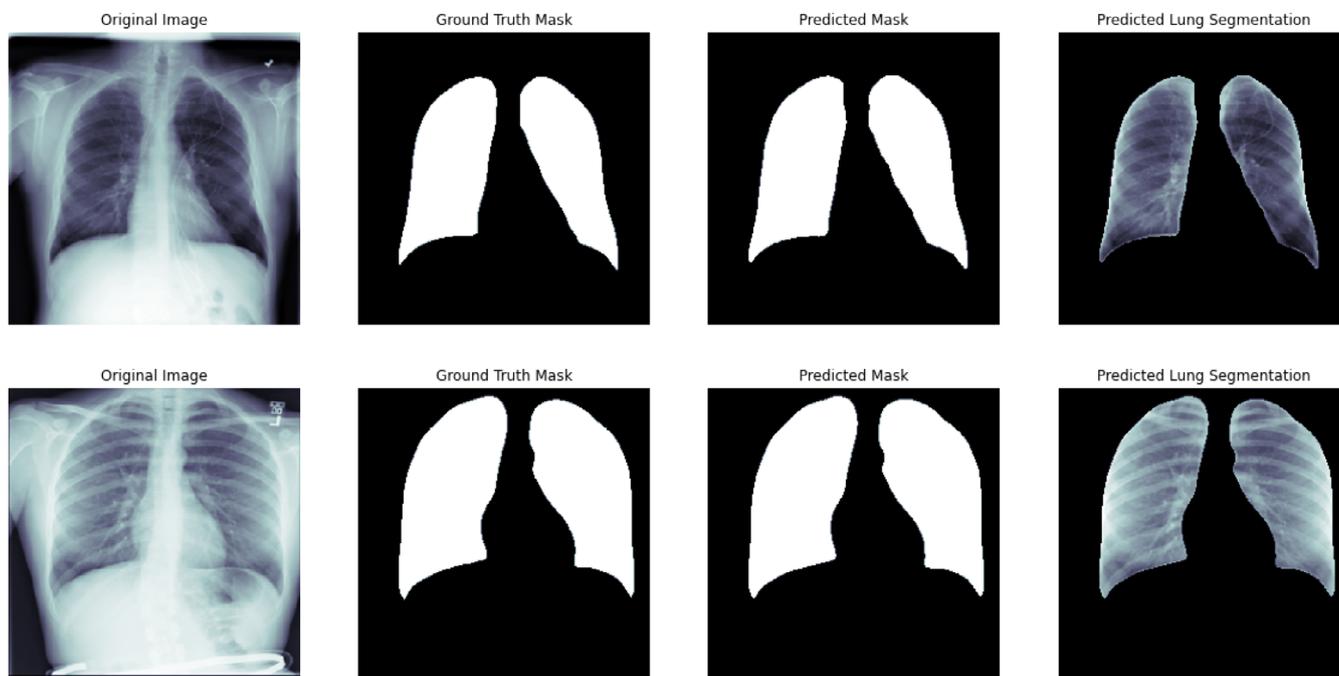
### 5.1 | Experimental setup

The Python 3.8.0 environment has been used to develop the experimental study on both plain and segmented lung X-ray images. The Tensorflow v2.8.2 platform, which offers pre-

TABLE 2 Summary of the proposed model 2’s architecture for three classes.

Layer	Output shape	Parameters
Input layer	224 × 224 × 3	0
Functional (DenseNet201)	7 × 7 × 1920	18,321,984
Functional (Xception)	7 × 7 × 2048	20,861,480
Functional (EfficientNetB7)	7 × 7 × 2560	64,097,687
Concatenate	7 × 7 × 6528	0
2D global average pooling	6528	0
Reshape	1 × 1 × 6528	0
Dense	1 × 1 × 816	5,327,664
Dense	1 × 1 × 6528	5,327,664
Multiply	7 × 7 × 6528	0
Flatten	319,872	0
Dense	64	20,471,872
Batch normalization	64	256
Dropout	64	0
Dense	3	195
Total parameters =		134,414,514
Trainable parameters =		31,133,235

built functions and sophisticated operations to make the effort of creating neural networks easier, was used for the experiment. The following tools and libraries were used in the study: Keras v2.8.0, OpenCV v4.1.2, Matplotlib v3.2.2, and Scikit-learn v1.1.2. The NVIDIA-SMI 460.32.03 Tesla T4 GPU with 460.32.03 Driver and 11.2 CUDA has been selected for open access by Google collaboratory. The optimal learning rate was maintained at 0.001 and continually tracked at each epoch over the loss function for the 20 epochs of training the classification networks. By dividing the train or validation data count by the batch size, the steps per epoch and validation step size were kept constant. Each time, 16 samples were run through the model in batches. The learning rate was specified to be decreased on the plateau while maintaining a value for the patience of



**FIGURE 12** CXR sample image (left), ground truth mask (first-middle), generated mask by the model (second-middle), segmented lung image (right).

**TABLE 3** Performance of the segmentation network.

Network	Accuracy	Dice loss	IoU	Dice coefficient
ResUNet	0.9829	0.0248	0.9451	0.9633
UNet	0.9793	0.0550	0.9124	0.9408

2, a verbosity of 1, a factor of 0.1, and a minimum learning rate of 0.00001. The evaluation was carried out by compiling using Adam as the optimizer. The performance measures were IoU, accuracy, precision, recall, and f1-score, and the loss functions: were dice-coefficient for segmentation and categorical cross-entropy for classification.

The following section demonstrates how well the segmentation model and classification model performed on the segmented lung images.

## 5.2 | Lung segmentation

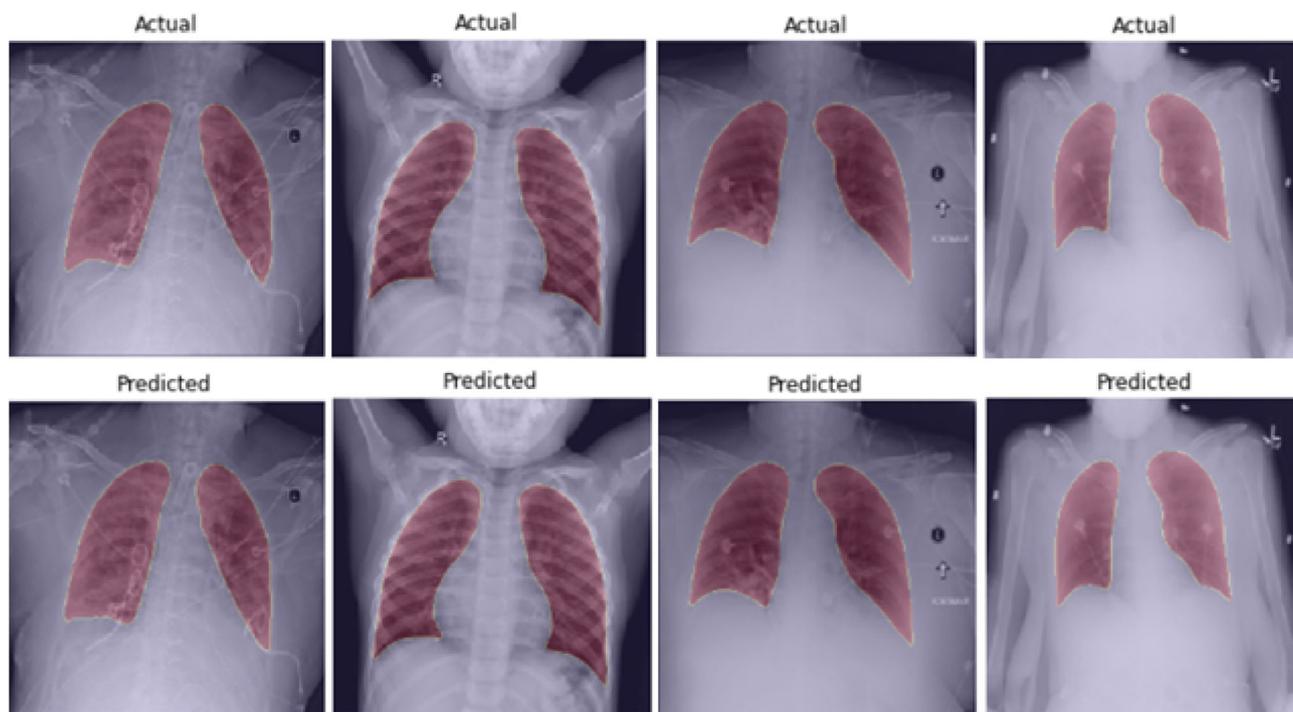
To assess the effectiveness of segmentation over the chest X-ray images, both the traditional U-net and the residual U-net model were trained, validated, and tested on the training, validation, and test data, respectively. However, the latter performed better than the former in terms of better accuracy as well as faster convergence [as seen in in Table 3]. A qualitative review proved the test images were appropriately segmented because the classification database's lack of ground truth masks prevented quantitative verification. The residual U-Net model, which was trained on the original chest X-ray images, can very reliably segment the lung sections of the X-ray images in the

classification database, as shown in Figure 12. A spectral visual representation of the segmented lungs for ground truth and predicted masks are shown in Figure 13. The segmentation model's total accuracy, loss, weighted IoU, and dice coefficient values are presented in Table 3.

## 5.3 | COVID recognition

As mentioned earlier, two separate experiments using two distinct proposed models have been run on our datasets, first looking at the three classes (COVID, normal, and viral pneumonia), then the four classes (COVID, lung opacity, normal, and viral pneumonia) taken as a whole. Along with that, the models' effectiveness on binary classes (COVID, normal) has also been observed. Table 4 displays the results of the classification performed on the original plain CXR images using each of the three pre-trained models (DenseNet201, EfficientNetB7, and Xception). The effectiveness of our suggested ensemble approaches outperformed that of any individual network. Tables 5 and 6 compare the effectiveness of the individual pre-trained models with the two ensemble techniques that have been tested for two and three classes on the segmented lung images, respectively.

The same comparative study is displayed for four classes in Table 7. It should be highlighted that the second approach model, which combined the models with a channel-wise attention module by saving the parameters of the pre-trained models after re-training on segmented images, performed better in all the circumstances. It should be mentioned here that the standard deviation values for accuracy and loss were also measured after re-running each model 5 times, where the standard deviations of all those runs are between 0.017 and 0.019.



**FIGURE 13** Spectral visualization of the segmented lungs for ground truth and predicted masks.

**TABLE 4** Performance of the three pre-trained models on plain CXR images.

Network	No. of classes	Accuracy	Loss
DenseNet201	2	0.9609	0.1140
	3	0.9210	0.1778
	4	0.8483	0.3954
EfficientNetB7	2	0.9362	0.1728
	3	0.9008	0.2157
	4	0.8430	0.3928
Xception	2	0.9279	0.1935
	3	0.8993	0.2410
	4	0.8320	0.4321

**TABLE 5** Performance of our models on two classes.

Network	Accuracy	Loss
DenseNet201	0.9457	0.1686
EfficientNetB7	0.9495	0.1657
Xception	0.9614	0.1473
Proposed model 1	0.9688	0.1372
Proposed model 2	0.9821	0.0560

The performance metrics for the first model in two, three, and four classes, respectively, are shown in Tables 8–10. The performance metrics for the second model on two, three, and four classes, respectively, are shown in Tables 11–13.

**TABLE 6** Performance of our models on three classes.

Network	Accuracy	Loss
DenseNet201	0.9311	0.1677
EfficientNetB7	0.9008	0.2410
Xception	0.9108	0.2157
Proposed model 1	0.906	0.2224
Proposed model 2	0.934	0.0791

**TABLE 7** Performance of our models on four classes.

Network	Accuracy	Loss
DenseNet201	0.9026	0.1943
EfficientNetB7	0.8849	0.2436
Xception	0.8892	0.2533
Proposed model 1	0.8589	0.2676
Proposed model 2	0.8906	0.2222

Table 14 displays the overall test accuracies for both models (proposed model 1 and proposed model 2) on unknown datasets.

For the two suggested models, the confusion matrices conducted on the test dataset for two, three, and four classes are shown in Figures 14 and 15, respectively.

In Figure 16, the correctly-classified and miss-classified COVID images with their corresponding true and predicted labels are depicted.

**TABLE 8** Categorization reports for two classes using proposed model 1.

	Precision	Recall	F1-score
COVID	0.95	0.88	0.91
Normal	0.93	0.97	0.95

**TABLE 9** Categorization reports for three classes using proposed model 1.

	Precision	Recall	F1-score
COVID	0.96	0.71	0.82
Normal	0.83	0.98	0.90
Viral pneumonia	0.99	0.89	0.94

**TABLE 10** Categorization reports for four classes using proposed model 1.

	Precision	Recall	F1-score
COVID	0.86	0.67	0.75
Lung opacity	0.80	0.89	0.84
Normal	0.87	0.92	0.89
Viral pneumonia	0.92	0.93	0.93

## 5.4 | Comparison with others and ablation study

We evaluated the effectiveness of the models that we suggested and then contrasted the findings with earlier research. By splitting it up into three parts, a comparison of our recommended approaches with prior attempts is presented in Tables 15–20, respectively. [21] used image-enhancement techniques to create better images while interfering with the properties of the original plain images in three classes, having one class different from ours. In order to achieve better results, they also used the conventional U-Net model, although this model is heavier than our

**TABLE 11** Categorization reports for two classes using proposed model 2.

	Precision	Recall	F1-score
COVID	0.99	0.92	0.95
Normal	0.95	1.00	0.97

**TABLE 12** Categorization reports for three classes using proposed model 2.

	Precision	Recall	F1-score
COVID	0.95	0.83	0.89
Normal	0.89	0.95	0.92
Viral pneumonia	0.90	0.89	0.89

**TABLE 13** Categorization reports for four classes using proposed model 2.

	Precision	Recall	F1-score
COVID	0.82	0.79	0.81
Lung opacity	0.86	0.88	0.87
Normal	0.92	0.93	0.92
Viral pneumonia	0.94	0.93	0.93

**TABLE 14** Overall independent test accuracy for two, three, and four classes on unknown data.

Network	No. of classes	Independent test accuracy
Proposed model 1	2	0.94
	3	0.88
	4	0.85
Proposed model 2	2	0.97
	2	0.91
	3	0.87

lightweight residual U-Net approach. [42] used a modified version of EfficientNet to do the three-class experiment with one class being different from ours, with marginal performance gain. [10] experimented with a novel customized model while taking into account a substantial dataset of high-quality images, producing results that are close to ours. [23]’s experiments with an ensemble CNN on a dataset with slightly more images than ours for three classes produced results that were comparable to ours, but their method suffers greatly when used to four-class experiments. [24] also used the COVID-Net approach, but the dataset contained fewer images, and his results were marginally lower than ours. Both [27] and [30] employed CovXNet, but the number of images in the dataset caused significant variances in their results. [38] employed DarkCovidNet, a variation of DarkNet, originally created to perform well with two classes, which performs better than our model in binary classification, but, however, our method outperformed it when the multi-class classification was considered. With their tried-and-true methods, the rest of the studies mentioned performed poorly than our method. In those above-mentioned methods, some have used data augmentation and additional preprocessing techniques, [10, 21, 23, 24, 30, 42], whereas the following papers [16, 25, 28, 29, 31, 32] haven’t used any preprocessing or enhancement techniques to interfere with the original image properties.

We assessed the effectiveness of DenseNet201, Xception, and EfficientNetB7 architectures for COVID-19 detection. We also proposed an attention-based ensemble model to combine their predictions, achieving enhanced performance compared to individual models. The findings demonstrate the importance of model ensembling and highlight avenues for further research in COVID-19 diagnosis using deep learning techniques. We conducted experiments using a dataset consisting of

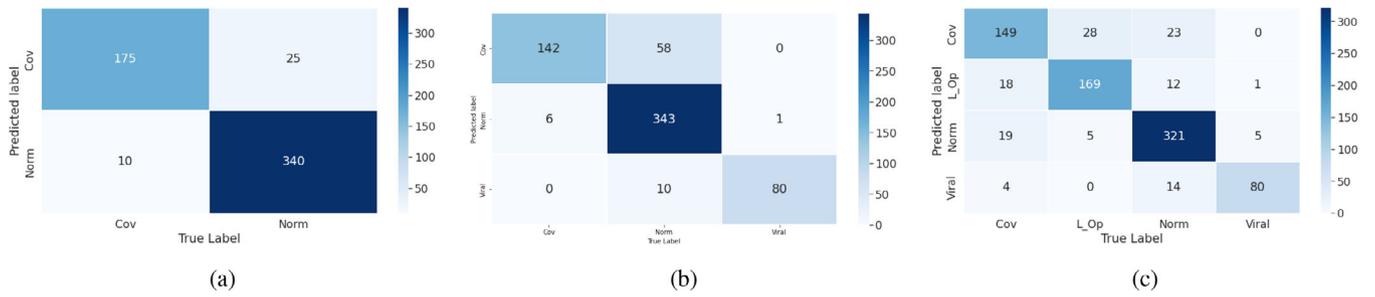


FIGURE 14 Confusion Matrices for (a)–(c) two, three, and four classes for model 1.

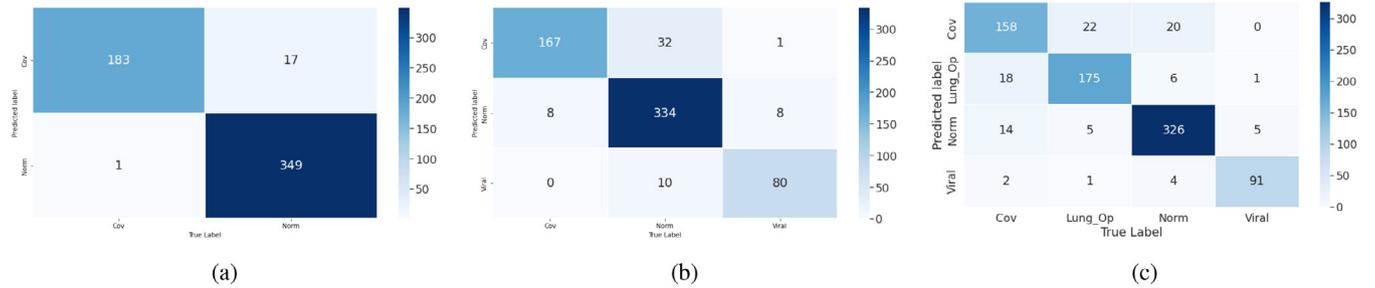


FIGURE 15 Confusion matrices for (a)–(c) two, three and four classes for model 2.

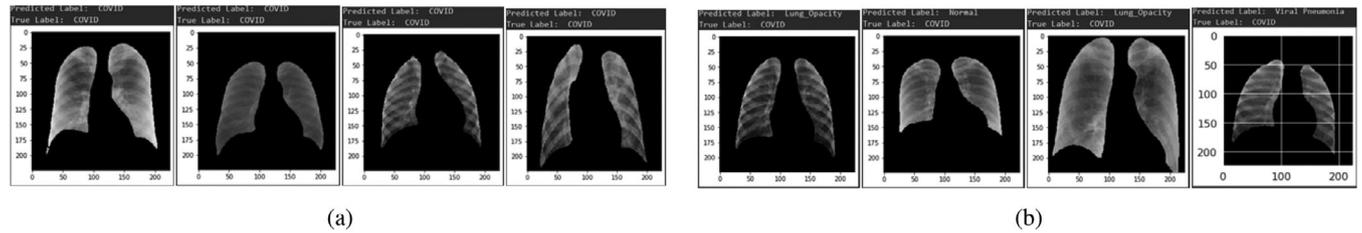


FIGURE 16 (a) Correctly-classified COVID images. (b) Misclassified COVID images.

TABLE 15 Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-1).

Study	Dataset count	Used method	Preprocessing and/or augmentation	Performance
[21]	18,479 (Normal - 8851, COVID - 3616, lung opacity - 6012)	U-Net and DenseNet201	Gamma correction	Accuracy: 0.9511
[10]	13,962 (COVID - 358, pneumonia - 5538, normal - 8066)	COVID-Net	Cropping, translation rotation, horizontal flip, zoom and intensity shift	Accuracy: 0.933
[16]	1127 (COVID - 127, no findings - 500, pneumonia - 500)	DarkCovidNet	–	Accuracy: 0.8702
[33]	618 (COVID - 219, viral pneumonia - 224, healthy - 175)	ResNet + location awareness	Resampling morphological operations (erosion, dilation, convex hull) clipping and flipping, mirroring	Accuracy: 0.867
[42]	13,569 (Normal, COVID, lung opacity)	Modified efficientNet	Intensity normalization, rotation, scaling horizontal flip	Accuracy: 0.939
[23]	15,478 (COVID, pneumonia, normal)	InceptionV3 NASNetLarge ensemble CNN	CLAHE with (blur, erosion, dilation, closing, hole filling) horizontal flip, zoom, shear	Accuracy: 0.90 Accuracy: 0.924 Accuracy: 0.931

**TABLE 16** Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-2).

[23]	5851 (Normal COVID bacterial pneumonia viral pneumonia)	MobileNet ResNet ensemble CNN	CLAHE with (blur, erosion, dilation, closing, hole filling) horizontal flip, zoom, shear	Accuracy: 0.768 Accuracy: 0.772 Accuracy: 0.815
[24]	1645 (COVID, lung ipacity, normal)	COVID-Net	–	Accuracy: 0.933
[29]	6570 (COVID - 714, healthy - 1583, bacterial pneumonia - 1860 viral pneumonia - 2413)	VGG16 + InceptionV3	–	Accuracy: 0.6250
[25]	1300 (COVID - 290, normal - 1203, bacterial pneumonia - 660 viral pneumonia - 131)	CoroNet	–	Accuracy: 0.896
[26]	1125 (healthy, COVID, pneumonia)	DenseNet169 + MobileNet and LightGBM	–	Accuracy: 0.9111
[27]	9915 (COVID - 3305, pneumonia - 3305, normal - 3305)	CovXNet	Hashing algorithm	Accuracy: 0.9202

**TABLE 17** Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-3).

[28]	1125 (control - 125, COVID - 500, pneumonia - 500)	COVID-19FcNet9	–	Accuracy: 0.8996
[30]	1248 (COVID - 215, pneumonia - 500, non-COVID - 533)	CovXNet	Rotation, axis shift, horizontal flip, scaling and shear transformation	Accuracy: 0.836
[32]	327 (normal - 152, COVID - 152, pneumonia - 50)	VGG16	–	Accuracy: 0.8410
[31]	400 (COVID - 100, pneumonia - 100, normal - 200)	VGG19	–	Precision: 0.83 sensitivity + F1 score: 0.80
[35]	15,085 (COVID - 180, pneumonia - 6054, normal - 8851)	Xception + ResNet50V2	Rotation, zoom, shifting horizontal/vertical flip	Accuracy: 0.914
[43]	2331 (COVID - 231, No finding - 1050, pneumonia - 1050)	CapsNet	Horizontal flip, width and height shift	Accuracy: 0.8422

**TABLE 18** Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-4).

[44]	5982 (COVID - 1765, normal, pneumonia)	ResNet101	–	Accuracy: 0.719
[34]	4352 (COVID - 1292, CAP - 1735, non pneumonia - 1325)	COVNet	–	Sensitivity: 0.90
[45]	— (COVID, normal, pneumonia)	Cascaded network	–	Sensitivity: 0.8936
[36]	52,000 (healthy - 26,000, COVID - 26,000)	MobileNetV2, VGG19, ResNet101	Rotation, horizontal flip, width and height shift	Accuracy: 0.98, Accuracy: 0.95, Accuracy: 0.95
[37]	835 (COVID - 420, normal - 415)	GAN + CNN	cGAN	Accuracy: 0.987
[38]	1127 (normal - 127, COVID - 500, pneumonia - 500)	DarkCovidNet	–	Accuracy: 0.8786 Precision: 0.9135

chest X-ray images categorized into four classes and employed transfer learning by initializing the pre-trained weights of the selected architectures. We proposed an ensemble model that combines the predictions of DenseNet201, Xception, and EfficientNetB7 using an attention mechanism. The attention mechanism assigns weights to each model's prediction based on its performance on the validation set. While designing the proposed architectures we encountered several settings by adding

or removing certain components, such as convolution blocks, utilizing different activation functions, or replacing dense layers and batch normalization layers with convolution and observed the strengths and weaknesses of each architecture in COVID-19 detection. The results shown in the Table 5–7 indicate each model's performance on its own as well as the advancement achieved through the proposed attention-based ensemble model. The ensemble model leverages the complementary

**TABLE 19** Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-5).

[38]	627 (COVID - 127, Normal - 500)	DarkCovidNet	–	Accuracy: 0.9897 Precision: 0.9846
Model 1 proposed	13,748 (Normal - 10192, COVID - 3616)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.9688
Model 1 proposed	15,153 (Normal - 10192, COVID - 3616, viral pneumonia - 1345)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.906
Model 1 proposed	21,165 (Normal - 10192, COVID - 3616, viral pneumonia - 1345, lung opacity - 6012)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.8589
Model 2 is being proposed	13,748 (Normal - 10192, COVID - 3616)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.9821

**TABLE 20** Comparison of the proposed COVID-19 diagnostic methods with other deep learning methods developed using radiology images (part-6).

Model 2 is being proposed	15,153 (normal - 10192, COVID - 3616, viral pneumonia - 1345)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.934
Model 2 is being proposed	21,165 (normal - 10192, COVID - 3616, viral pneumonia - 1345, lung opacity - 6012)	Residual U-Net & ensemble(Xception) DenseNet201, EfficientNetB7) + attention	N/A	Accuracy: 0.8906

features learned by different architectures, leading to improved classification accuracy.

## 5.5 | Discussion

In this study, we have developed a deep convolutional neural network architecture for automatic COVID-19 detection in a computer-aided diagnosis model utilizing segmented chest X-ray images. After utilizing the ResUNet model to segment the lung portion from X-ray images, we used an ensemble of three pre-trained models (DenseNet201, Xception, and EfficientNetB7) to achieve enhanced prediction accuracy. CXR images were chosen over more traditional means because they are less expensive, simpler to use, and faster. The ResUNet Model was chosen for its quicker memory access and shorter processing times; reasonable performance with fewer parameters. We have a large dataset with a lot of images, thus implementing class imbalance techniques can be computationally expensive and resource-intensive, requiring a lot more processing power and memory space. Moreover, important information in the image or the quality of the image may occasionally be distorted as a result of geometric transformations. The intended meaning may change or details may be lost as a result of this distortion. Additionally, the findings could be erroneous depending on the quality of the transformation or the parameters that were utilized. Because of this, we have disregarded geometrical transformations and image enhancement. Transfer learning has received praise for reducing resource consumption and providing effective training. The slight class imbalance challenges associated with our dataset are addressed by using ensemble technology and fine-tuned transfer learning models, which contribute to enhancing overall performance, particularly for minority classes. To solve the problem of vanishing gradients and limit the number of parameters, DenseNet-201 is

employed. To achieve greater accuracy and smaller footprints, EfficientNetB7 is used. The reason Xception was picked is that it classifies things thoroughly and into numerous categories. The attention channel is employed to enhance the interrelations between channels and concentrate on specific extracted features. The use of ensemble techniques increases the reliability and efficiency of forecasts. To the best of our knowledge, our proposed second model performed better than many previous cutting-edge experimental types of research, with higher accuracy readings of 98.21% for two classes, 93.4% for three classes, and 89.06% for four classes, respectively. We discovered that our two proposed ensemble models had outperformed some other existing models after examining Tables 15–20. Some studies' performance might be superior to ours because those studies' image types differed in some way from ours. In the majority of the investigations, there were various numbers of datasets. The researchers who used the same dataset as ours may have performed better since they used manual preprocessing approaches, and some of them have taken the accuracy of their validation data as their final accuracy rather than doing experiments on unseen image data. The dataset's owner, the paper [21], published a better performance on classes other than ours and the usage of image-enhancing techniques prior to classification might have resulted in improved accuracy. Other than that, our proposed model worked effectively even without any image enhancement or preprocessing methods.

## 6 | CONCLUSION

Controlling infectious diseases and pandemic circumstances like the current COVID-19 requires a quick diagnosis technique. The COVID-19 virus spread like wildfire and had an impact on the global economy. A good substitute to outperform the results of the current investigations is AI-based technology.

Apart from what we contributed to the research, there are usually certain limitations that we were unable to overcome. Firstly, our dataset has not been used in clinical trials. Field data collection and accurate annotation performed in person may increase efficiency. As there are currently only four different image class types in the dataset, we must evaluate the suggested approach using chest radiography images of other sorts of diseases in order to generalize it. An important obstacle to performing classification was an unbalanced dataset count. Moreover, identification of the true class label is a difficult operation to do because of the CXR images' ground glass opacity. With regard to accuracy and computational time, our deep learning-based approach for COVID diagnosis still has a lot of room for growth. Applying ensemble to segmentation may be the focus of our future study. More hybrid deep model exploration is possible. The model's accuracy and dependability can be increased by more hyperparameter testing and adjustment on unobserved data. The image's hot regions can be emphasized using Grad-CAM-based visualization. A thorough experiment can be performed with some experienced professionals. In order for a model to be considered robust, it must be able to generalize across a variety of populations, including individuals with varied genetic backgrounds and geographic locations. More and more training on data from multiple places is required for the model to operate more effectively. It's crucial to remember that doctors and other medical professionals frequently use multiple diagnostic tools to find COVID-19. To make a more reliable diagnosis, a combination of techniques is used instead, including clinical evaluation, laboratory tests, and imaging. In addition to other diagnostic techniques like RT-PCR testing and CT imaging, our proposed model should be viewed as a complementary tool. Combining various diagnostic modalities can potentially increase precision and reliability in general. It is anticipated that the impressive accuracy of the proposed segmentation and classification identification technique noted in the Section 5, in the COVID-19 worldwide epidemic will establish a mechanism for COVID-19 patients to lower the burden and viral proliferation associated with COVID-19 clinical diagnostics. We can infer that this study has theoretical significance for developing methods to provide robust classifiers with strong reporting ability as well as usefulness in real-world applications.

## AUTHOR CONTRIBUTIONS

**Abu Sayeed:** Conceptualization; formal analysis; investigation; project administration; resources; supervision. **Nasif Osman Khansur:** Conceptualization; formal analysis; investigation; methodology; software; visualization; writing—original draft. **Azmair Yakin Srizon:** Conceptualization; formal analysis; investigation; software; validation; visualization; writing—original draft. **Md. Farukuzzaman Faruk:** Investigation; methodology; software; validation; visualization; writing—review and editing. **Salem A. Alyami:** Investigation; methodology; software; validation; writing—review and editing. **AKM Azad:** Investigation; methodology; software; validation; writing—review and editing. **Mohammad Ali Moni:** Conceptualization; formal analysis; funding acquisition; project

administration; resources; supervision; validation; writing—original draft.

## ACKNOWLEDGEMENTS

This work was supported and funded by the Deanship of Scientific Research at Imam Mohammad Ibn Saud Islamic University (IMSIU) (grant number IMSIU-RPP2023003).

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in "COVQU" at <https://doi.org/10.1016/j.combiomed.2021.104319>, reference number.

## ORCID

*Mohammad Ali Moni*  <https://orcid.org/0000-0003-0756-1006>

## REFERENCES

1. Wu, F., Zhao, S., Yu, B., Chen, Y.-M., Wang, W., Song, Z.-G., Hu, Y., Tao, Z.-W., Tian, J.-H., Pei, Y.-Y., et al.: A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265–269 (2020)
2. Huang, C., Wang, Y., Li, X., Ren, L., Zhao, J., Hu, Y., Zhang, L., Fan, G., Xu, J., Gu, X., et al.: Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395, 497–506 (2020)
3. WHO: Coronavirus (2020). <https://www.who.int/health-topics/coronavirus>. Accessed 12 Oct 2022
4. WHO: Pneumonia of unknown cause – China (2022).
5. Uddin, I.: A brief introduction to COVID-19 (corona virus disease). *J. Med. Healthcare*. 2(3), 1–2 (2020)
6. Worldometers: Covid live - coronavirus statistics - Worldometer (2022). <https://www.worldometers.info/coronavirus/>. Accessed 12 Oct 2022
7. Rubin, G.D., Ryerson, C.J., Haramati, L.B., Sverzellati, N., Kanne, J.P., Raoof, S., Schluger, N.W., Volpi, A., Yim, J.-J., Martin, I.B.K., et al.: The role of chest imaging in patient management during the COVID-19 pandemic: a multinational consensus statement from the fleischner society. *Radiology* 296(1), 172–180 (2020)
8. Pray, I.W., Ford, L., Cole, D., Lee, C., Bigouette, J.P., Abedi, G.R., Bushman, D., Delahoy, M.J., Currie, D., Cherney, B., et al.: Performance of an antigen-based test for asymptomatic and symptomatic SARS-CoV-2 testing at two university campuses-Wisconsin, september–october 2020. *Morb. Mortal. Wkly. Rep.* 69(51–52), 1642 (2021)
9. Park, J., Stevens, R., Freer, R., Soneji, N., Jones, N.: The Accuracy of Chest CT in the Diagnosis of COVID-19: An Umbrella Review. Oxford University, Oxford (2022)
10. Wang, L., Lin, Z.Q., Wong, A.: COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci. Rep.* 10(1), 1–12 (2020)
11. Zhao, W., Zhong, Z., Xie, X., Yu, Q., Liu, J., et al.: Relation between chest CT findings and clinical conditions of coronavirus disease (COVID-19) pneumonia: a multicenter study. *AJR Am. J. Roentgenol.* 214(5), 1072–1077 (2020)
12. Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z.A., Zhang, N., Diao, K., Lin, B., Zhu, X., Li, K., et al.: Chest CT findings in coronavirus disease-19 (COVID-19): relationship to duration of infection. *Radiology* 295(3), 200463 (2020)
13. The Global Economy: Doctors per 1,000 people by country, around the world. [https://www.theglobaleconomy.com/rankings/doctors\\_per\\_1000\\_people/](https://www.theglobaleconomy.com/rankings/doctors_per_1000_people/) (2020). Accessed 12 Oct 2022
14. Alzubaidi, L., Al-Amidie, M., Al-Asadi, A., Humaidi, A.J., Al-Shamma, O., Fadhel, M.A., Zhang, J., Santamaría, J., Duan, Y.: Novel transfer learning

- approach for medical imaging with limited labeled data. *Cancers* 13(7), 1590 (2021)
15. Adaloglou, N.: Transfer learning in medical imaging: classification and segmentation (2020). <https://theaisummer.com/medical-imaging-transfer-learning/>. Accessed 12 Oct 2022
  16. Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* 121, 103792 (2020)
  17. Narin, A., Kaya, C., Pamuk, Z.: Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* 24(3), 1207–1220 (2021)
  18. Bhattacharyya, A., Bhaik, D., Kumar, S., Thakur, P., Sharma, R., Pachori, R.B.: A deep learning based approach for automatic detection of COVID-19 cases using chest X-ray images. *Biomed. Signal Process. Control* 71, 103182 (2022)
  19. Wang, W., Li, Y., Li, J., Zhang, P., Wang, X.: Detecting COVID-19 in chest X-ray images via MCFff-Net. *Comput. Intell. Neurosci.* 2021, 3604900 (2021)
  20. Keidar, D., Yaron, D., Goldstein, E., Shachar, Y., Blass, A., Charbinsky, L., Aharony, I., Lifshitz, L., Lumelsky, D., Neeman, Z., et al.: COVID-19 classification of X-ray images using deep neural networks. *Eur. Radiol.* 31(12), 9654–9663 (2021)
  21. Rahman, T., Khandakar, A., Qiblawey, Y., Tahir, A., Kiranyaz, S., Kashem, S.B.A., Islam, M.T., Al Maadeed, S., Zughaier, S.M., Khan, M.S., et al.: Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. *Comput. Biol. Med.* 132, 104319 (2021)
  22. Hussain, E., Hasan, M., Rahman, M.A., Lee, I., Tamanna, T., Parvez, M.Z.: CoroDet: a deep learning based classification for COVID-19 detection using chest X-ray images. *Chaos Solitons Fractals* 142, 110495 (2021)
  23. Lin, T.-C., Lee, H.-C.: COVID-19 chest radiography images analysis based on integration of image preprocess, guided grad-cam, machine learning and risk management. In: *Proceedings of the 4th International Conference on Medical and Health Informatics*, pp. 281–288. ACM, New York (2020)
  24. Teixeira, L.O., Pereira, R.M., Bertolini, D., Oliveira, L.S., Nanni, L., Cavalcanti, G.D.C., Costa, Y.M.G.: Impact of lung segmentation on the diagnosis and explanation of COVID-19 in chest X-ray images. *Sensors* 21(21), 7116 (2021)
  25. Khan, A.I., Shah, J.L., Bhat, M.M.: Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. *Comput. Methods Programs Biomed.* 196, 105581 (2020)
  26. Nasiri, H., Kheyroddin, G., Dorrigiv, M., Esmacili, M., Nafchi, A.R., Ghorbani, M.H., Zarkesh-Ha, P.: Classification of COVID-19 in chest X-ray images using fusion of deep features and LightGBM. *arXiv:2206.04548* (2022)
  27. Sampén, D., Lavarello, R.: Comparison of deep learning architectures for COVID-19 diagnosis using chest X-ray images. In: Mello-Thoms, C.R., Taylor-Phillips, S. (ed.) *Medical Imaging 2022: Image Perception, Observer Performance, and Technology Assessment*, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 12035, pp. 120350Z. SPIE, Bellingham, WA (2022)
  28. Barua, P.D., Muhammad Gowdh, N.F., Rahmat, K., Ramli, N., Ng, W.L., Chan, W.Y., Kuluozturk, M., Dogan, S., Baygin, M., Yaman, O., et al.: Automatic COVID-19 detection using exemplar hybrid deep features with X-ray images. *Int. J. Environ. Res. Public Health* 18(15), 8052 (2021)
  29. Kaushik, B., Koundal, D., Goel, N., Zaguia, A., Belay, A., Turabieh, H.: Computational intelligence-based method for automated identification of COVID-19 and pneumonia by utilizing CXR scans. *Comput. Intell. Neurosci.* 2022, 7124199 (2022)
  30. Nishio, M., Noguchi, S., Matsuo, H., Murakami, T.: Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: combination of data augmentation methods. *Sci. Rep.* 10(1), 1–6 (2020)
  31. Horry, M.J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., Shukla, N.: X-ray image based COVID-19 detection using pre-trained deep learning models (2020)
  32. Moutounet-Cartan, P.G.B.: Deep convolutional neural networks to diagnose COVID-19 and other pneumonia diseases from posteroanterior chest X-rays. *arXiv:2005.00845* (2020)
  33. Xu, X., Jiang, X., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Ni, Q., Chen, Y., Su, J., et al.: A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* 6(10), 1122–1129 (2020)
  34. Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., et al.: Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* 296(2), E65–E71 (2020)
  35. Rahimzadeh, M., Attar, A.: A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inf. Med. Unlocked* 19, 100360 (2020)
  36. Akter, S., Shamrat, F.M.J.M., Chakraborty, S., Karim, A., Azam, S.: COVID-19 detection using deep learning algorithm on chest X-ray images. *Biology* 10(11), 1174 (2021)
  37. Karakanis, S., Leontidis, G.: Lightweight deep learning models for detecting COVID-19 from chest X-ray images. *Comput. Biol. Med.* 130, 104181 (2021)
  38. Nair, R., Vishwakarma, S., Soni, M., Patel, T., Joshi, S.: Detection of COVID-19 cases through X-ray images using hybrid deep neural network. *World J. Eng.* 19(1), 33–39 (2021)
  39. Rahman, T., Chowdhury, M., Khandakar, A.: COVID-19 radiography database (2021). <https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>. Accessed 11 Oct 2022
  40. Chowdhury, M.E.H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z.B., Islam, K.R., Khan, M.S., Iqbal, A., Al Emadi, N., et al.: Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access* 8, 132665–132676 (2020)
  41. Lakhani, P.: The importance of image resolution in building deep learning models for medical imaging. *Radiol. Artif. Intell.* 2(1), e190177 (2020)
  42. Luz, E., Silva, P., Silva, R., Silva, L., Guimarães, J., Miozzo, G., Moreira, G., Menotti, D.: Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. *Res. Biomed. Eng.* 38(1), 149–162 (2022)
  43. Toraman, S., Alakus, T.B., Turkoglu, I.: Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* 140, 110122 (2020)
  44. Che Azemin, M.Z., Hassan, R., Mohd Tamrin, M.I., Md Ali, M.A.: COVID-19 deep learning prediction model using publicly available radiologist-adjudicated chest X-ray images as training data: preliminary findings. *Int. J. Biomed. Imaging* 2020, 8828855 (2020)
  45. Yeh, C.-F., Cheng, H.-T., Wei, A., Chen, H.-M., Kuo, P.-C., Liu, K.-C., Ko, M.-C., Chen, R.-J., Lee, P.-C., Chuang, J.-H., et al.: A cascaded learning strategy for robust COVID-19 pneumonia chest X-ray screening. *arXiv:2004.12786* (2020)
  46. Pan, S.J., Yang, Q.: A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* 22(10), 1345–1359 (2010)
  47. Christodoulidis, S., Anthimopoulos, M., Ebner, L., Christe, A., Mougiakakou, S.: Multisource transfer learning with convolutional neural networks for lung pattern analysis. *IEEE J. Biomed. Health Inform.* 21(1), 76–84 (2016)
  48. Yang, H., Mei, S., Song, K., Tao, B., Yin, Z.: Transfer-learning-based online Mura defect classification. *IEEE Trans. Semicond. Manuf.* 31(1), 116–123 (2017)
  49. Akcay, S., Kundegorski, M.E., Willcocks, C.G., Breckon, T.P.: Using deep convolutional neural network architectures for object classification and detection within X-ray baggage security imagery. *IEEE Trans. Inf. Forensics Secur.* 13(9), 2203–2215 (2018)
  50. Zhang, Z., Liu, Q., Wang, Y.: Road extraction by deep residual U-Net. *IEEE Geosci. Remote Sens. Lett.* 15(5), 749–753 (2018)
  51. Huang, G., Liu, Z., Pleiss, G., Van Der Maaten, L., Weinberger, K.: Convolutional networks with dense connectivity. *IEEE Trans. Pattern Anal. Mach. Intell.* 44(12), 8704–8716 (2019)

52. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: International Conference on Machine Learning, pp. 6105–6114. Microtome Publishing, Brookline, MA (2019)
53. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1251–1258. IEEE, Piscataway, NJ (2017)
54. Sun, G., Hu, J., Shen, L.: Squeeze-and-excitation networks. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 7132–7141. IEEE, Piscataway, NJ (2018)

**How to cite this article:** Sayeed, A., Khansur, N.O., Srizon, A.Y., Faruk, M.F., Alyami, S.A., Azad, A., Moni, M.A.: An effective screening of COVID-19 pneumonia by employing chest X-ray segmentation and attention-based ensembled classification. *IET Image Process.* 18, 2400–2416 (2024).  
<https://doi.org/10.1049/ipr2.13106>